VOLUME 54
NUMBER 2

WHOLE NO. 244
1942

Psychological Monographs

The Nature and Measurement of Musical Meanings

By

K. BRANTLEY WATSON, PH.D.
Duke University

Published by
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Publications Office
NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS

THE NATURE AND MEASUREMENT OF MUSICAL MEANINGS

CHAPTER I

INTRODUCTION

GENERAL STATEMENT OF THE PROBLEM

THE specific purpose of this investigation is twofold: (1) to determine the general nature of musical meanings[1] at different age-levels and (2) to devise an instrument whereby ability to discriminate between such meanings can be measured.

A few investigators have reported that different types of music can be interpreted with a fair degree of consistency by adults. There are no data, however, on the nature of children's responses, nor has an attempt been made to measure discriminative ability objectively at any level.

Authorities agree that at all age-levels music appreciation is dependent upon "an understanding" of music. It is therefore important to know the precise nature of musical meanings as interpreted not only by adults but by children as well. Likewise, a measure of ability to discriminate between such meanings would be valuable as a measure of one aspect of music appreciation.

EXPLORATORY INVESTIGATION

Since there have been no comprehensive experimental data reported on the nature of musical meanings for children, an exploratory investigation was conducted for the purpose of determining the most suitable approach to the problem.

Several orchestral recordings which seemed to the writer to represent different types of music were selected. Short excerpts of from one to two minutes in length were played for eight individual subjects at each of the fourth-, fifth-, sixth-, eighth-, and twelfth-grade levels and at the adult level. The subjects were asked to "think out loud" while listening to the music and were, at times, prompted by suggestions from the experimenter, such as: "What does this music mean to you?", "How does the music make you feel?", etc. The listeners were asked, further, to state what characteristics of the music determined its meaning. *Verbatim* reports were taken of the subjects' answers and voluntary responses.

The data indicated that a number of different types of meanings may be attributed to music. Those most frequently observed were as follows:[2]

1. *Objective* (correlated directly with the physical stimulus itself—simple and complex perceptions). Example: "That is a high note." "Now the melody is outstanding."

[1] No attempt is made in this study to define the concept of "meaning" epistemologically. Since the investigation is primarily experimental, the term will be used broadly to include all types of response to music. The writer does not wish to imply that the problem of the fundamental nature of meanings is unimportant. The question, however, is not of particular significance in this study.

[2] Further data were obtained on the frequency of mention of the various types of meanings for different age-levels, growth in recognition of the different types, etc. These data are omitted, however, since they are not directly related to the major problem.

2. *Imaginal* (creating imaginary situations which might be represented by the music). Example: "It's like a party, and everyone is running around playing."
3. *Associational* (interpreting the music in terms of some concrete experience). "It's a band playing." "It sounds like thunder in a storm.
4. *Abstract* (characterizing the music in terms of a mood). Example: "It's happy." "It's very exciting."
5. *Subjective* (interpreting the music in terms of the way it affects the subject himself). Example: "It stirs me all up inside." "It makes me depressed."

In order to determine the consistency with which subjects at different age-levels could classify music according to meanings, a list of simple adjectives (abstract meanings) representing various characterizations of music was formulated. In addition, lists of imaginal, associational, and subjective types of meanings were devised. New subjects at each age-level were asked to indicate orally the items in the various lists which best fitted the different selections as they were played. The subjects were also asked to rate each selection in terms of certain *musical attributes* (pitch, volume, tempo, rhythm, melody, harmony, etc.). The meaning of each of these terms and the terms used in the lists of meanings was carefully explained before the ratings were made. Questions pertaining to the subjects' answers were asked at the conclusion of each selection in order to obtain more detailed information concerning the nature of the musical response.

The following tentative observations were suggested from an analysis of the data:

1. Certain types of music can be consistently characterized by certain meanings, this consistency increasing with the higher grade-levels.

2. There is a consistent relationship at all levels between certain patterns of musical attributes and certain meanings.[3] It is possible to describe these musical meanings in terms of musical attributes at all age-levels, this consistency increasing with the higher levels.

3. There are significant individual differences in the ability to discriminate between musical meanings of the kind measured.

4. The observations made above are most striking in connection with abstract meanings. This type of meaning is most amenable to the testing technique employed, due, probably, to the more intangible subjective factors which predominantly influence the other types of derived meanings.[4]

SPECIFIC STATEMENT OF THE PROBLEM

The observations stated above served as a basis for formulating the following specific aims of the present study.

1. To devise a test of ability to discriminate between abstract musical meanings.

2. To determine what specific abstract meanings may be represented musically.

3. To ascertain the musical characteristics of the different abstract meanings.

4. To find the consistency with which music can be interpreted according to meaning at different grade-levels.

5. To analyze the general nature of growth in ability to discriminate between abstract musical meanings.

[3] Expressed in terms of "types" of meanings, "musical attributes" refers to "objective meanings." The statement may therefore be read: "There is a consistent relationship at all levels between certain patterns of primary meanings and the other types of (derived) meanings."

[4] In addition to furnishing the most productive experimental approach to the problem of musical meanings, abstract meanings are also recognized by musicians as being more significant musically than any of the other types of derived meanings. For these reasons the analysis of musical meanings will be made in this study solely from the standpoint of "abstract" meanings and their relationship to "objective" meanings.

The present investigation was not instigated for the purpose of proving a psychological theory, or even of evolving one, in regard to the nature of musical meanings or discriminative ability. The purpose was, rather, of more general nature; namely, to describe the broader characteristics of musical meanings[5] as they could be analyzed from the measurement technique employed. As such, the study serves as a basis for a more detailed investigation, being conducted at the present time, of the specific problems suggested.

[5] Hereafter, in this report, the term "meaning" will refer specifically to "abstract" meaning of the kind described above.

Chapter II

Construction and Standardization of the Test of Ability to Discriminate Between Abstract Meanings

The general procedure followed in constructing the test of ability to discriminate between musical meanings was to devise a check-list of appropriate adjectives (abstract meanings) and to select musical recordings representative of each of the list items. Subjects could then indicate on the check-list the adjectives which best represented each selection as it was played. In previous studies of musical meanings no attempt has been made to determine experimentally what adjectives are appropriate for representing different types of music or, consequently, what musical selections best represent the abstract classifications. This experimental fallacy no doubt accounts in part for the divergence of opinion among other investigators as to the nature and degree of constancy of musical meanings. In the present investigation the most appropriate adjectives to be used in the check-list were determined experimentally in a series of preliminary tests, and appropriate corresponding musical representations were likewise objectively selected. The final form of the test was then standardized on the basis of the judgments of expert musicians.

First Preliminary Test

The writer listened to approximately three hundred recordings of orchestral music by recognized composers and selected fifty which seemed to represent the greatest possible variety of meanings. Records were selected from which short excerpts of from one to two minutes in length and representing complete passages (that is, having musical beginnings and endings) could be taken.

In addition, a check-list of sixty adjectives representing as wide a variety of meanings as possible was formulated. Included in the list were adjectives suggested by subjects in the exploratory investigation. Provision was also made on the check-list for rating each selection in terms of its musical attributes; namely, pitch, volume, tempo, consonance, dynamics, rhythm, melody, and harmony.

The fifty excerpts were played for a group of ten musicians who were asked to indicate on the check-list, for each selection played, all adjectives which were appropriate for describing the music. The adjective which best described the music was to be indicated by a *1*, the next by a 2, and all others selected were to be checked. If a selection could not be classified by any abstract meaning, it was to be checked as "unclassified."

Each selection was also to be rated on the basis of its musical attributes. Rhythm, melody, and harmony were to be rated *1*, 2, or *3* on the basis of the predominance of the particular attribute. Pitch, volume, tempo, consonance, dynamics, and regularity of rhythm were to be rated on a scale of three points, *1* representing one extreme (for example, very low in pitch), *3* the other extreme (for example, very high in pitch), and 2 "medium."

From the data obtained the following tabulations were made:[6]

[6] Wherever it was possible to do so without greatly subtracting from the significance of the

1. *The relationship between different adjectives.* The frequency with which each adjective was checked in combination with every other adjective was determined. Frequencies were converted into per cents of total frequency of mention. Thus, *glad,* one of the adjectives in the check-list, was checked sixty times. *Happy* was checked forty times along with *glad.* The per cent relationship is therefore 40/60, or 67%. *Mischievous* was checked fourteen times with *glad.* The per cent relationship is therefore 14/60, or 23%. Similar relationships were determined between each adjective and every other adjective.

2. *The musical characteristics of each adjective.* The average pitch, volume, tempo, consonance, dynamics, and regularity of rhythm, was determined for all selections classified by individual subjects under a given adjective. Thus, for example, of the sixty times *happy* was checked by individual musicians the selection so classified was given a rating in pitch of *3* twenty-five times, of *2* thirty-three times, and of *1* two times. The average pitch for *happy* is therefore 2.4.

Similar data were obtained on the relative predominance of rhythm, melody, and harmony for all adjectives.

3. *The frequency with which each adjective was mentioned for each of the selections played.* By allowing three points for an adjective given first place, two points for a second place, and one point for a third place (or check), the relative weights of each adjective for each record were determined. Thus, for the first record played, *glad* was given a weighting of twenty points, *happy* twenty-one points, *merry* fifteen points, *joyful* ten points, *exuberant* five points, *playful* six points, and so on.

4. *The average musical characteristics of each selection played, regardless of its adjective-classification.* Averages were determined as for 2 above, except that tabulations were made for individual records rather than for adjectives.

First Revision of the Check-List

The per cent relationship of the different adjectives and the musical characteristics of the adjectives (*1* and 2 above) served as a basis for revising the check-list. The most closely related adjectives (those with the highest per cent relationship) were combined into single check-items. Some adjectives in the original list were eliminated because of infrequency of mention. Other adjectives were eliminated because of too broad a meaning; that is, they were fairly closely related to many other adjectives but not significantly related to any one. A few adjectives which were not included in the original list, but which were added frequently by the subjects voluntarily, were included in the revised list.

In a few instances it was impossible from the data on the relationship of adjectives to determine whether a certain adjective should be included in the revised list. In such instances the average musical characteristics of the different meanings served as a further basis for selection. For example, one particular adjective was not closely related to any of the others. That it did, however, represent a distinct type of meaning could be seen from the fact that its musical characteristics were different from those of any of the others. It was therefore included in the revised list as a separate item.

The revised check-list, representing the

---

context, tabular and other statistical data have been eliminated from this condensation of the original study. Complete data can be found in the original manuscript, a Ph.D. dissertation, on file in the Duke University Library, Durham, North Carolina.

combination of appropriate adjectives, consisted of eighteen different adjective-groups, and a nineteenth item designated as "unclassified." The list of adjective-groups is given below.

LIST OF REVISED ADJECTIVE GROUPS

Happy, glad
Playful, mischievous
Amusing, funny
Exuberant, very happy
Stirring, impetuous
Exciting
Sensational, stormy, very exciting
Dignified, sedate, stately
Pompous, majestic, kingly
Peaceful, dreamy, tranquil
Serious, sober, solemn, reverent
Pleading, yearning, passionate
Somber, gloomy
Sad, mournful, doleful
Desolate, tragic
Strange, mysterious, wierd
Foreboding, ominous, afraid
Grotesque, wild
Unclassified

With regard to the rating of musical attributes, it was suggested by several subjects that a wider choice would give more precise data. In the revised test form, therefore, the range of choice was extended from three points to five. Thus, for example, the possible choices for pitch were: *1* very low, *2* low, *3* medium, *4* high, *5* very high.

Likewise, it was pointed out that certain records could not be rated by a single number. For example, the pitch might change from very low to very high during the course of the selection. Allowance was made, therefore, for rating each selection as to its range in the particular attribute. Thus, if the range of pitch in a particular selection was from very low to very high, both *1* and *5* would be checked.

Selection of Records

The records which had been most consistently characterized by the individual adjective-groups (as revised) were selected for use in a second revision of the test. In a few instances, where it was impossible to determine which of several records was the best, the one was chosen whose musical characteristics more nearly approximated those of the adjective-group for which it was selected.

In addition to the records thus selected, enough new ones were added to provide three selections representative of each adjective-group in the revised test form. These additional records were chosen by the writer, who attempted to find selections with musical characteristics similar to those representing the adjective-group for which the record was chosen.

SECOND PRELIMINARY TEST

The fifty-seven selections were played for twelve new musicians (two sittings) who indicated their responses on the revised form. Instructions were as follows:

I will play a number of musical recordings. After the group of adjectives which best describes each selection write a number *1;* after your second choice write a number *2;* write a number *3* after any other groups which you think are at all appropriate. If the music has some meaning other than can be described by the words listed, give your own description. If the music has no particular meaning at all, place a check after "unclassified."

Likewise, rate each selection according to its musical attributes by drawing a circle around the point in the scale which you think is appropriate. If there is a noticeable change in the characteristic, you may indicate the range of the change by drawing circles around two of the points. For example, if a selection seems to you to change in pitch from very low to very high, you would draw a circle around both the *1* and the *5* in the scale of pitch.

Rhythm, melody, and harmony are to be rated on the basis of their predominance.

Indicate your first choice by a *1*, your second choice by a 2.

Second Revision of the Check-List and Selection of Records to be Used in the Test

The following tabulations were made from the data obtained:[7]

1. *The frequency with which each adjective-group was mentioned for each of the records played.* By allowing three points for an adjective-group mentioned as first choice for a given record, two points for a second choice, and one point for a third choice, adjective weightings were determined for each selection.

2. *The musical characteristics of each adjective-group (average pitch, volume, etc.).* Average musical characteristics were determined as described for the first preliminary test.

3. *The musical characteristics of each selection (average pitch, volume, etc.).* Musical characteristics of the different selections were also determined as described for the first preliminary test.

These data were used for a further revision of the check-list and for a selection of the most appropriate records.

It was found that a reliable discrimination could be made between fifteen of the eighteen adjective-groups. These fifteen groups are those included in the final test form, shown on page 10.

From the fifty-seven records used in this second preliminary test, the two which were most consistently classified by the twelve musicians for each adjective-group were selected for final standardization. There were, then, thirty records finally chosen, two for each of the fifteen revised adjective-groups. The list of records, with a description of the excerpts, is shown in Table 1.

[7] Complete data may be found in the original manuscript.

FINAL STANDARDIZATION OF THE TEST

The thirty records finally selected were played for twenty expert musicians (tested individually or in small groups)[8] who indicated their responses on the revised test form (see page 43). Pitch, volume, tempo, sound, dynamics, and rhythm were rated by checking the appropriate point or points on the scale of *1* to *5*.[9] Under rhythm, melody, and harmony the element noticed most was checked. In case two of the elements or all three were equally noticeable, two or all three could be checked. The choice of meaning for each selection was indicated by placing the numbers of appriate adjective-groups in the spaces provided for the first, second, and third choices.

The subjects were all musicians of recognized training and ability. All were instructors in college departments of music. Eight were instructors in music theory and composition, three were directors of choral groups, four were directors of symphony orchestras, two were instructors in piano, one was an instructor in violin, and two were instructors in public school music.

Derivation of Weighted Scores for the Test

Weighted scores for the test were derived from the experts' test data. Although provision was made on the test blank for indicating three possible meanings for each selection played, only the first and second choices were used in deriving the weights.

Four different methods of scoring the

[8] Only two of the twenty musicians had participated in the preliminary testing, and they only in the first preliminary test.

[9] The provision on the test form for rating musical attributes (pitch, volume, etc.) was made solely for experimental purposes. The test itself consists only of rating each record in terms of its meaning.

TABLE 1

Description of the Thirty Records* Chosen to Represent the Different Adjective-Groups

Group	Selection	Description of Excerpt	Time Limit
1A†	Gounod: Dance of the Nubians	$\frac{3}{16}$ in. in on record‡	1:28
1B	Handel: Menuet	Beginning of record	1:01
2A	Mendelssohn: Scherzo	Beginning of record	1:20
2B	Gluck-Gevaert: Tambourin	Beginning of record	:58
3A	Moussorgsky: Ballet of the Unhatched Chicks	$\frac{1}{8}$ inch in on record	1:14
3B	Haydn: Toy Symphony, Part I	Beginning of record	:59
4A	Haydn: Symphony in G Major, Finale	Last of record	:45
4B	Beethoven: Fifth Symphony, Fourth Movement	$\frac{1}{2}$ inch in on record	1:10
5A	Beethoven: Overture to Egmont, Part 2	Last of record	1:15
5B	Weber: Overture to der Freischutz	Last of record	1:00
6A	Tschaikowsky: Francesca da Rimini, Part 4	$1\frac{1}{2}$ in. from end	1:20
6B	Berlioz: Rakoczy March	$1\frac{3}{4}$ in. from end	1:15
7A	Handel: Alcina Suite Overture	Beginning of record	:58
7B	Brahms: Symphony No. 1 in C Minor, Fourth Movement (concl.)	First of record	1:00
8A	Rimsky-Korsokow: Cortege	Beginning of record	1:02
8B	Lulli: March from Thesee-	First of record	:48
9A	Debussy: En Bateau	First of record	1:20
9B	Beethoven: Sixth Symphony, Second Movement	First of record	1:23
10A	Bach: Komm Susser Tod	First of record	1:15
10B	Mendelssohn: Nocturne	First of record	1:08
11A	Schonberg: Verklarte Nachte, Part 2	$\frac{3}{8}$ inch in on record	1:23
11B	Tschaikowsky: Sixth Symphony, First Movement (cont.)	$1\frac{7}{8}$ inch in on record	1:30
12A	Grieg: Peer Gynt Suite, Ingred's Lament	$\frac{1}{4}$ inch in on record	1:05
12B	Rachmaninoff: The Isle of the Dead, Fourth Movement	Last of last record	1:27
13A	Rachmaninoff: The Isle of the Dead, Part 2	In $1\frac{1}{2}$ in. on record	1:37
13B	Tschaikowsky: Sixth Symphony, Fourth Movement	In 1 inch on record	1:13
14A	Berlioz: Fantastic Symphony, Dream of a Sabbat	Beginning of record	1:00
14B	Stravinsky: Le Sacre du Printemps, Dance of the Earth	End of record	1:20
15A	Bach: Chaconne (cont.)	Beginning of record	1:17
15B	Scarlatti-Taussig: Pastorale	Beginning of record	:55

* The writer is greatly indebted to the R.C.A. Victor Co. for furnishing all the records listed in the table except *En Bateau*, by Debussy, and *Pastorale*, by Scarlatti-Taussig.

† All selections listed *A* (for example, 1A, 2A, etc.) constitute Form A of the test; all selections listed *B* constitute Form B. In each form of the test, then, there are 15 records, one for each of the fifteen adjective-groups.

‡ The excerpt was begun $\frac{3}{16}$ inch in from the beginning of the record. White ink was used to indicate the place where the excerpt began and ended.

test (when used with subjects subsequently tested) were actually tried out: (1) allowing credit for only "correct" answers (answers given most frequently by the experts) for first, second, or third choice; (2) allowing weighted credit for only the first choice for each selection; (3) allowing weighted credit for all three

choices; and (4) the method to be described, that of allowing weighted credit for the first two choices. Method four was selected since it provided the highest reliability coefficient when the test was administered to 100 college freshmen.

The weighting for first choice was determined by allowing 2 points for an expert's first choice and 1 point for a second choice. Thus, if 14 experts classified "Dance of the Nubians" as *Group 1* for their first choice and 5 classified this selection as *Group 1* for second choice, the weighting of *Group 1* as a first choice for this record would be 14 times 2, plus 5 times 1, or 33. The weighting allowed for selecting this group as second choice was the total number of times this group was mentioned, either as a first or second choice. Thus, for "Dance of the Nubians," the weighting of *Group 1* as a second choice would be 14 (number of first choices) plus 5 (number of second choices), or 19 (the total number of times *Group 1* was mentioned for this record).

The weightings thus derived for each adjective-group for all records are shown in Table 2. In each instance the first row in the table following an adjective-group represents the weightings of that adjective when selected as a first choice, the second row representing the weightings of the same adjective when selected as a second choice.

The table is read as follows: *Happy* is given a weighting of 33 as a first choice for record 1 ("Dance of the Nubians"; see Table 1 for a list of records) and a weighting of 19 as a second choice for this record; for Record 1a (alternate record chosen to represent *Group 1*) *happy* is given a weighting of 29 for first choice and 16 for second choice. It will be noticed from the data that although *happy* is given the highest weighting for Records 1 and 1a it is also given some weighting for Records 2, 2a, 3, 3a, and so on.

The weightings shown in Table 2 were next converted into per cents of the highest possible weight. Forty is the highest possible weight a single adjective-group could be given for a single record (if all twenty experts selected it as first choice). These per cents were then converted into probable error differences above and below the mean possible score (20). Each of the probable error scores was multiplied by 2 to allow whole-point credit for the lowest scores. The value of −3 P. E.'s was designated as the scale zero, since this point allowed some credit (1 point) for an adjective-group having a weighting of 2 points but allowed none for a group having a weighting of only 1 point. It was believed that a weighting of only 1 point might in many instances represent only a chance subjective factor.

Derivation of P. E. scores is illustrated by the following example. An adjective weighting of 37 (the highest weighting attributed to any adjective-group for any record by the experts) is 92.5% of the total possible weighting (40), or 42.5% above the mean possible weighting (20). The value of 42.5% corresponds to 2.13 P. E.'s. Scaled from −3 P. E. this would be 5.13 P. E. (3 plus 2.13). This figure multiplied by 2 gives a rough weighting of 10. Similarly, an adjective weighting of 33 gives a probable error score of 9, as does a weighting of 32; a weighting of 31 gives a probable error score of 8; and so on.

The points to be allowed for different answers on the test were derived by substituting these probable error scores (weights) for the gross weighting of adjective-groups given in Table 2. In Table 3 are shown the number of points (probable error scores, weights) to be allowed for various answers for each record

TABLE 2

Adjective Weightings of the Different Records, First and Second Choices (Expert Classification)

Adjective-Group	Number of Record																											
	1	1a	2	2a	3	3a	4	4a	5	5a	6	6a	7	7a	8	8a	9	9a	10	10a	11	11a	12	12a	13	13a	14	14a	15	15a
Happy	33*	29	13	16	5	13	12	15	4	3				1	3	5	4												11	13
	19	16	9	11	4	9	10	10	4	2				1	3	3	3												6	10
Mischievous	5	12	30	28	23	17	9	12	1																		1			
	4	9	17	16	16	12	6	9	1																		1			
Amusing			7	2	28	20	1	1																						
			6	2	17	12	1	1																						
Very happy	6	7	6	14	4	7	28	23	9	6		1			1	6													7	
	5	5	5	11	3	5	15	14	8	5		1			1	4													5	
Exciting				4		2	9	9	28	26	22	15			11	6											4	9	2	
				3		1	6	6	16	15	17	12			7	4											4	9	2	
*Very exciting								1	9	12	33	26															2	9		
								1	5	9	19	16															2	8		
Dignified	4	7				1				4			31	27	16	15	8	9	3	6			1				2		7	10
	4	6				1				4	1		18	16	12	12	6	7	2	5			1				1		5	7
Kingly									9	9		10	19	19	30	23				3									4	
									6	5		3	15	15	18	14				3									3	
Peaceful	12	5											10	13			32	25	13	17	5	3	2		1					14
	8	4											7	8			18	15	10	10	3	3	1		1					9
Serious													10	13			7	8	30	29	9	3	17	9	2	1	1			
													7	8			5	6	19	17	6	3	10	7	2	1	1			
Pleading																	2	6		1	30	32			1	17				
																	2	4		1	18	18			1	12				
Sad																			14	4	7	15	29	29	16	15	1			
																			9	4	6	11	18	18	13	11	1			
Tragic											5	2									9	7	10	20	29	27	4	14		
											4	2									7	5	9	15	15	16	4	10		
Mysterious												4											1	2	11		37	28		
												4											1	2	8		19	15		
Unclassified							1									5	7	12									8		29	23
							1									3	6	8									7		19	14

* The top figure after each adjective-group refers to the weighting for first choice; the bottom figure (in this case *19*) refers to weighting for second choice.

played in the test. Thus, if on a later testing a subject selects *happy, glad* as a first choice for Record *1*, he is allowed 9 points (weighting of 33, Table 2, gives a score of 9 points when converted into a probable error score). If *happy, glad* is mentioned as a second choice for this record, 6 points are allowed (weighting of 19, Table 2, gives a probable error score of 6). If *playful, mischievous* is given as first choice for this record, 3 points are allowed. No credit is given for mentioning *amusing, funny,* and so on. The total number of points so obtained by a subject for first and second choices for all records is the subject's score on the test. Thus a subject is given most credit for an answer which was given most frequently by the experts, but is also given some credit for other answers, the amount of credit depending upon how many experts had given the same answer.

Validity of the Test

The validity of the test, as a test of discrimination of musical meanings, is dependent in part upon the expertness of the musicians' judgments. In addition to the fact that all the experts were musicians of recognized ability, further evidence of the validity of their judgments can be seen in a comparison of their analyses with those of nationally recognized music critics. In every instance the experts' analyses agreed with program notes, analyses, and interpretations of the recognized critics, and, where they were available, with the interpretations offered by the composers themselves.[10] A further evidence of the value of the expert judgments is the consistency with which each record was analyzed.

There is some question as to the validity of the test at lower age-levels, due to possible vocabulary difficulties.[11] It may well be that for children the test is one of vocabulary as well as one of musical discrimination. Children at the fourth-, fifth-, and sixth-grade levels were tested individually to determine the suitability of the vocabulary.[12] Considerable evidence of vocabulary difficulty was found in the fourth grade. In the fifth grade, when the terms were explained prior to the testing, little apparent difficulty was encountered. That is, when the writer asked the children to explain why a particular selection was happy, or exciting, or sad, etc., the answers were consistent with the meaning of the adjective itself, even though the correct musical meaning of the selection might be quite different.

Objective evidence of the suitability of the vocabulary was obtained from a comparison of the abstract and imaginal meanings, as found in the exploratory investigation. As was described in Chapter I, subjects listened to records and checked on prepared lists of abstract and imaginal meanings the items in each list which best described the music. In no instances were there any inconsistencies between imaginal and abstract meanings above the fourth grade.[13] For example, "like a little boy playing tricks on people" was checked along with *playful, mischievous,* or *funny, amusing* but was never checked along with *dignified* or *tragic* or *serious.*

On the check-lists used in the preliminary investigation *very, very happy;*

[10] Perhaps the best single reference for purposes of comparison is: Charles O'Connell, *The Victor Book of the Symphony* (New York: Simon and Schuster, Inc., 1934).

[11] The test is to be used later for testing children at the sixth-grade level.

[12] This testing was a continuation of the testing reported in the exploratory investigation.

[13] Tabular data may be found in the original manuscript.

TABLE 3

Points Allowed for Different Answers on the Test
(First and Second Choices)

Adjective-Groups	Number of Record																											
	1	1a	2	2a	3	3a	4	4a	5	5a	6	6a	7	7a	8	8a	9	9a	10	10a	11	11a	12	12a	13	13a	14	14a	15	15a
Happy	9*	8	5	5	3	5	4	5	2	2					2	3	2												4	5
	6	5	4	4	2	4	4	4	2	1					2	2	2												3	4
Mischievous	3	4	8	8	7	5	4	4																						
	2	4	5	5	5	4	3	4																						
Amusing			3	1	8	6																								
			3	1	5	4																								
Very happy	3	3	3	5	2	3	8	7	4	3						3													3	
	3	3	3	4	2	3	5	5	3	3						2													3	
Exciting			2			1	4	4	8	7	6	5			4	3											2	4	1	
			2			1	3	3	5	5	5	4			3	2											2	3	1	
Very exciting									4	4	9	7															1	4		
									3	4	6	5															1	3		
Dignified	2	3								2			8	7	5	5	3	4	2	3									3	4
	2	3								2			6	5	4	4	3	3	1	3									3	3
Kingly									4	4		4	6	6	8	7				2									2	
									3	3		3	5	5	6	5				2									2	
Peaceful	4	3											4	5			9	7	5	5	3	2								5
	3	2											3	3			6	5	4	4	2	2								
Serious													4	5			3	3	8	8	4	2	5	4	1					
													3	3			3	3	6	5	3	2	4	3	1					
Pleading																	1	3			8	9				5				
																	1	2			6	6				4				
Sad																			5	2	3	5	8	8	5	5				
																			4	2	3	4	6	6	4	4				
Tragic											3	1									4	3	4	6	8	8	2	5		
											2	1									3	3	4	5	5	5	2	4		
Mysterious												3						4						1	4		10	8		
												2						3						1	3		6	5		
Unclassified	•															3	3										3		8	7
																3	3										3		6	5

* The top figure after each adjective-group refers to the weighting for first choice; the bottom figure (in this case *6*) refers to weighting for second choice.

very, very exciting; kingly, majestic, pompous; and *pleading, yearning, passionate* were not included. It is assumed that if the subjects know the meaning of *exciting* and *happy* they will also know the meaning of the superlatives of these adjectives. This leaves only *pleading, etc.* and *kingly, etc.,* about which no information was obtained. In the later individual testing these adjectives were included among the possible choices, and appropriate records were included among those used. The choices of imaginal responses were (1) "like wanting something so much and you can't get it" and (2) "like a king and all his royal princes." Of ten subjects, eight chose *pleading* for (1) and two chose *sad.* The two who chose *sad* selected *pleading* as a second choice. All ten subjects chose *kingly* for (2).[14]

It is recognized that the evidence cited above is not entirely conclusive. It would not be expected that imaginal situations would have the same meaning for all subjects, even if there were no question of differences in understanding vocabulary. The most conclusive type of evidence, however, cannot be reported. This was obtained from asking the children individually, prior to the actual testing for musical meanings, to tell what the words meant. In some instances the children could not read the word but knew what it meant after it had been pronounced for them. A few minutes' drill on reading the words aloud removed this difficulty in every instance. The only adjective-groups where any difficulty of explaining meanings was encountered were *Groups 7 (dignified, stately, sedate, proud), 10 (serious, sober, solemn, reverent), 11 (pleading, yearning, passionate),* and *13 (tragic, desolate).* All of these were readily understood when explained by the experimenter, with the possible exception of *Group 7.* The writer could not determine if all the children understood the meaning of this group of adjectives, even though their responses were never wholly inconsistent with the meaning. Further evidence will be cited later, however, which is more conclusive from an objective standpoint.

The writer is aware of the many factors which bear on this problem of vocabulary meanings. In the first place, meanings cannot be classified dichotomously into categories of "understood" and "not understood." Understanding must always be evaluated on the basis of degrees of comprehension. When the statement is made in this study that subjects understand the terms used in the test, the reference is to verbalization of meanings, which involves giving a correct definition of a term or an appropriate synonym. Without becoming further involved in a theoretical discussion of the nature of meanings, it can be said that children in the fifth and sixth grades can at least give an appropriate verbal meaning for the terms used in the test.

In order to minimize the vocabulary factor, subsequent testing was not conducted at the fifth-grade level but was begun at the sixth-grade—and as a further precaution, each of the adjectives used was carefully explained before each testing in the lower levels. The musical terms (rhythm, melody, etc.) were explained and illustrated at all levels.

Reliability of the Test

Correlations obtained between two forms of the test (100 college students) were as follows for the different methods of scoring:

[14] It is possible that in some instances the selection of "kingly" for "like a king and all his royal princes" is due to the occurrence of "king" in both forms, rather than to a fundamental understanding of the term. Such, however, is improbable.

1. Weighted credit for correct answer ..... .60
2. Weighted credit for only first choice ..... .70
3. Weighted credit for all three choices ..... .83
4. Weighted credit for the first two choices ..... .85[15]

Using the probable error scores, a reliability coefficient of .90 was obtained (same 100 college students) between two halves (Forms A and B) of the test. This figure, stepped up by the Spearman-Brown "prophecy" formula, would be .95 for the entire test of thirty selections.

Correlation between the two halves of the test (given on successive days) was .80 for fifty sixth-grade pupils. This figure stepped up would be .89.

The retest reliability was determined for 49 college students (interval of three weeks). The coefficient was .89.

[15] As was explained earlier, the method of allowing weighted credit for the first two choices was selected as the best method of scoring since it yielded the highest reliability coefficient. All coefficients were obtained from different methods of scoring the same 100 papers.

Chapter III

General Nature of Musical Meanings for Expert Musicians

The data from the standardization procedure indicate that music can be classified consistently according to at least fifteen different types of abstract meanings by expert musicians. For each record played there was some meaning which was noticeably predominant. The fact that records could be selected from the analyses of one group of musicians and then be consistently interpreted by another group of musicians is in itself conclusive evidence in support of this contention.

The purpose of the present chapter is to analyze descriptively the general nature of such meanings.

Musical Characteristics of the Different Meanings

The general nature of musical meanings for expert musicians was analyzed first from the standpoint of the musical characteristics of the different meanings. Three types of analyses were made: (1) the average musical characteristics of the adjective-groups; (2) the relative predominance of rhythm, melody, and harmony for the different groups; and (3) the average ranges of the different attributes for each group.[16]

Average Musical Characteristics of Adjective-Groups

Average musical characteristics for a given group were computed by finding the average characteristics of all selections to which this particular meaning was attributed by individual experts. If, for example, ten musicians gave a rating of *3* in pitch to all selections which they characterized as *happy* and ten gave a rating of *4* in pitch to all selections characterized as *happy*, the average pitch for *happy* would be 3.5. Similar averages for each of the musical attributes were determined for each adjective-group. Only responses where a single point on the scale of attributes was mentioned (for example, pitch: low 1 (2) 3 4 5 high) were included in the calculation of averages.

Relative Predominance of Rhythm, Melody, and Harmony for the Different Groups

Weightings for the relative predominance of rhythm, melody, and harmony for the different adjective-groups were computed as follows: If a selection characterized as *glad* was rated outstanding in rhythm, with melody as the next outstanding attribute, rhythm was given a rating of two points for *glad* and melody was given a rating of one point. The total number of points given by all subjects to a given attribute for a given adjective was the weighting of that attribute for that particular adjective-group.

Average Ranges of the Different Attributes for Each Group

Frequently subjects preferred to indicate a range of choice (for example, pitch: low 1 (2) 3 (4) 5 high) rather than

[16] Specific data on the average musical characteristics of the different groups, on the per cent predominance of rhythm, melody, and harmony, and on the average ranges for musical attributes can be found, together with a detailed analysis of each, in the original manuscript.

TABLE 4

Complete Musical Characteristics of the Different Adjective-Groups

Adjective-Group	Averages and Standard Deviations of Musical Attributes												Order Predominance			Range of Musical Attributes
	Pitch	σ	Vol.	σ	Tempo	σ	Sound	σ	Dyn.	σ	Rhythm	σ	R	M	H	
Happy	3.4	.58	3.0	.57	3.2	.56	1.9	.56	2.5	.51	2.1	.66	1	2	3	Small in all
Mischievous	3.7	.70	3.0	.77	4.3	.76	2.4	.78	3.6	.59	2.9	.80	1	2	3	Large in P and R; small in others
Amusing	3.3	.87	3.1	.65	3.2	.83	3.0	.80	3.7	.51	3.6	.78	1	2	2	Large in P, V, R; small in others
Very happy	3.6	.65	3.7	.62	4.2	.48	2.0	.78	3.5	.61	2.7	.72	1	2	3	Large in P and V; small in others
Exciting	3.9	.82	4.1	.58	4.0	.53	3.0	.83	3.2	.89	2.8	.80	1	2	2	Medium in all
Very exciting	4.2	.83	4.6	.67	4.3	.56	3.5	.86	3.4	.80	3.1	.87	1	3	2	Medium in P, large in V; small in others
Dignified	2.8	.53	3.0	.46	2.4	.61	1.7	.51	1.9	.70	1.6	.63	1	1	2	Small in all
Kingly	3.2	.72	3.5	.68	2.9	.77	2.3	.71	2.5	.79	2.0	.58	1	2	2	Small in all
Peaceful	3.2	.83	2.4	.58	2.2	.69	2.3	.73	1.6	.47	1.6	.62	3	1	2	Small in all
Serious	2.2	.49	2.5	.57	1.9	.61	1.9	.58	2.0	.58	1.7	.59	3	1	2	Small in all
Pleading	3.9	.80	3.5	.82	3.2	.78	3.0	.72	3.5	.82	3.0	.80	3	1	2	Large in P and V; medium in others
Sad	1.7	.61	2.7	.71	1.6	.52	3.3	.69	1.6	.62	2.4	.78	3	2	1	Small in all
Tragic	2.5	.87	3.5	.68	2.6	.63	3.8	.59	3.0	.65	3.4	.69	3	3	1	Large in P and V; small in others
Mysterious	3.0	.90	2.8	.73	2.8	.52	4.3	.61	3.4	.70	4.0	.80	3	3	1	Large in P and V; medium in others
Unclassified	3.2	.63	3.1	.52	3.2	.61	2.7	.66	3.0	.59	2.8	.79	2	1	2	Small or medium in all

a single point on the scale. The average of these ranges was determined for each adjective-group. The procedure used was as follows (for average pitch range): In thirty-seven instances in which a selection was designated as *happy,* a single point on the scale of pitch was indicated. Range of pitch in these cases would be zero. In seven instances a range of one point (2 to 3, or 3 to 4, etc.) was indicated. In two instances a range of two points was indicated (1 to 3, or 2 to 4, etc.). In one instance a range of three points was indicated (1 to 4, or 2 to 5). The average of all these ranges (including zero ranges) is .40. It may be said, then, that the average pitch range for *Group 1 (happy, glad)* is .40 points. Similarly, average ranges were determined for each of the attributes for all groups.

SUMMARY OF MUSICAL CHARACTERISTICS OF THE DIFFERENT GROUPS

The data to which the preceding sections refer indicate that there are consistent differences between the musical characteristics of all of the fifteen adjective-groups. In Table 4 the complete musical description of each of the groups is given. Read as follows across the first row of the table: The adjective-group *happy, glad (happy)* is of above average pitch (3.0 being average),[17] with a standard deviation in pitch of .58, of average volume, of slightly above average tempo, of below average sound (dissonance), of below average dynamics (irregularity), and of below average rhythm (irregularity); it is outstanding in rhythm (melody second); and the average range of all musical attributes is relatively small.

RELIABILITY OF THE DIFFERENCES BETWEEN MUSICAL CHARACTERISTICS OF THE DIFFERENT ADJECTIVE-GROUPS

From the average musical characteristics of each group, with their standard deviations, the reliability of the difference between groups can be determined. In Table 5 is presented a sample of the computations which were made.[18] The six rows in the table following each adjective-group show the instances in which reliable differences, determined on the basis of critical ratio, occur for pitch (P, row 1), volume (V, row 2), tempo (T, row 3), sound (S, row 4), dynamics (D, row 5), and regularity of rhythm (R, row 6). Lack of reliable difference is indicated by an asterisk. The table is read as follows: *Group 1 (happy, glad)* is not reliably different in pitch (*) from *Groups 2, 3, 4, 8, 9, 11, 14,* and *15;* it is less in pitch (not so high, −P) than *Groups 5* and *6;* it is higher in pitch (P) than *Groups 10, 12,* and *13;* and so on for the other attributes.

Thus, from the sample table it is seen that *Group 1* is reliably different from *Group 2* in tempo, sound, dynamics, and regularity of rhythm; it is reliably different from *Group 3* in sound, dynamics, and regularity of rhythm; and so on. Similar data were obtained on the relationship between the musical characteristics of each group and every other group. There is a reliable difference be-

[17] Care must be taken, in interpreting the averages of musical characteristics, to allow for differences in standard deviations. The average pitch of a given group, for example, may be 3, but this does not indicate that all records selected as representing this group were of average pitch. It may well indicate a distribution of choices throughout the entire range (1 to 5). Such a situation would be represented by a large standard deviation of choices around the mean, and would indicate that pitch is not a determining factor for this group—that the group could be represented by selections of any pitch.

[18] Complete data will be found in the original manuscript.

TABLE 5

Reliable Differences Between the Musical Characteristics of Each Group and Every Other Group (Sample)

Adjective-Group	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	*	*	*	−P	−P	P	*	*	P	*	P	P	*	*
(1)	*	*	−V	−V	−V	*	−V	V	V	*	*	−V	*	*
Happy,	−T	*	−T	−T	−T	T	*	T	T	*	T	T	*	*
Glad	−S	−S	*	−S	−S	*	*	*	*	−S	−S	−S	−S	−S
	−D	−D	−D	−D	−D	D	*	D	D	−D	D	*	−D	*
	−R	−R	*	*	−R	*	*	*	*	−R	*	−R	−R	−R
		*	*	*	*	P	*	P	P	*	P	P	*	*
(2)		*	−V	−V	−V	*	−V	V	V	*	*	*	*	*
Mischievous,		T	*	*	*	*	T	T	T	T	T	T	T	T
Playful		*	*	*	−S	S	*	*	*	*	−S	−S	−S	*
		*	*	*	*	D	D	D	D	*	D	*	*	D
		*	*	*	*	R	R	R	R	*	R	*	*	R
			*	−P	−P	*	*	*	P	−P	P	P	*	*
(3)			−V	−V	−V	*	*	V	V	*	*	*	*	*
Amusing,			−T	−T	−T	T	*	T	T	*	T	*	*	*
Funny			S	*	*	S	*	*	S	*	*	−S	−S	*
			*	*	*	D	D	D	D	*	D	*	*	*
			R	*	*	R	R	R	R	*	R	*	*	R
				*	−P	P	*	*	P	*	P	P	*	P
(4)				*	−V	V	*	V	V	*	V	*	V	V
Exuberant,				*	*	T	T	T	T	T	T	T	T	T
Very happy				−S	−S	*	*	*	*	−S	−S	−S	−S	*
				*	*	D	D	D	D	*	D	*	*	*
				*	*	R	R	R	R	*	*	−R	−R	*
					*	P	P	*	P	*	P	P	*	P
(5)					−V	V	*	V	V	V	V	*	V	V
Exciting,					*	T	T	T	T	T	T	T	T	T
Stirring					*	S	*	*	S	*	*	−S	−S	*
					*	D	D	D	D	*	D	*	*	*
					*	R	*	R	R	*	*	*	−R	*
						P	P	P	P	*	P	P	P	P
(6)						V	V	V	V	V	V	V	V	V
Very exciting,						T	T	T	T	T	T	T	T	T
Sensational,						S	S	S	S	*	*	*	*	S
Stormy						D	D	D	D	*	D	*	*	*
						R	R	R	R	R	R	*	−R	*

tween all of the fifteen groups in at least one musical characteristic. If predominance of rhythm, melody, and harmony and range of musical attributes (not included in the calculation of reliable differences) are considered, the differences between groups would be even more striking.

INTERRELATIONSHIP OF MEANINGS

The general nature of musical meanings was further analyzed from the standpoint of the interrelation of adjective-groups. The number of times each of the groups was mentioned as a second choice in combination with each of the other groups (when mentioned as a first choice) was determined. Thus, *Group 1* was mentioned fifty-four times as a first choice. Of these times, *Group 2* was mentioned sixteen times as a second choice. *Group 3* six times, and so on. The per cent relationship of second choices to first choices was determined. For ex-

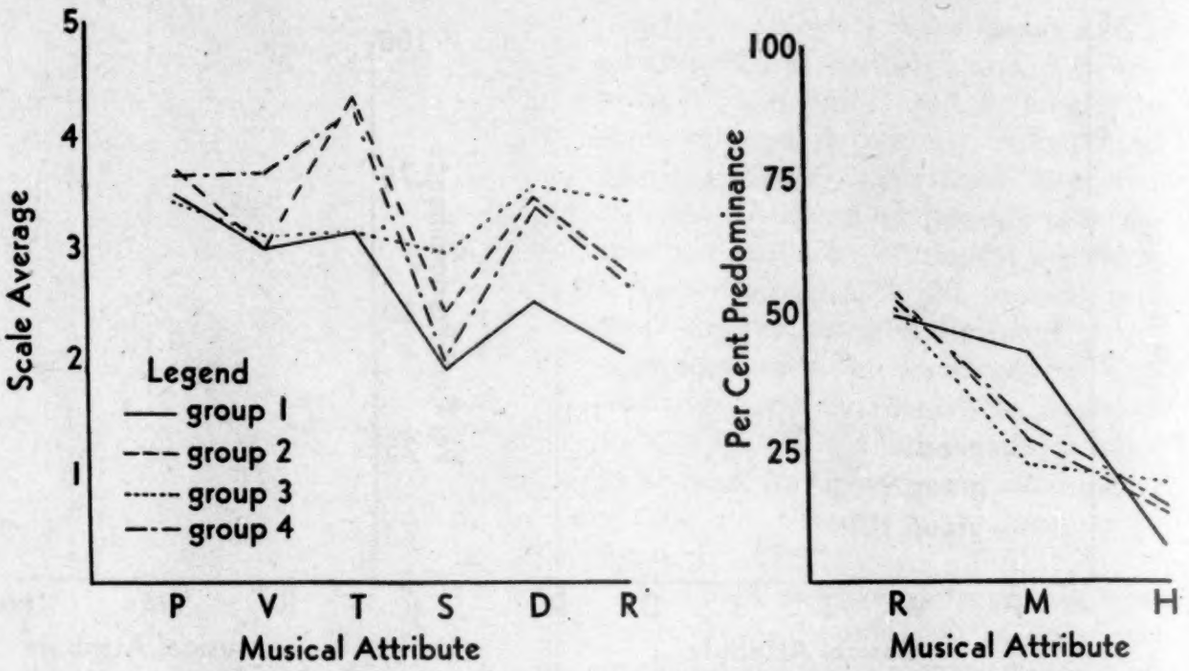

FIG. 1. Musical profiles for Groups 1 (happy), 2 (playful), 3 (amusing), and 4 (very happy)

ample, from the figures given above, *Group* 2 was mention 29.6% times (of the total possible times) as a second choice when *Group* 1 was mentioned as a first choice. Per cent relationship was thus determined between all groups. The data are presented in Table 6. The first row of the table is read as follows: Of all the times that *Group 1 (happy, glad)* was mentioned as a first choice, *Group 2 (playful, mischievous)* was mentioned 29.6% times as a second choice, *Group 3 (amusing, funny)* was mentioned 11.1% times, and so on. The per cent relation-

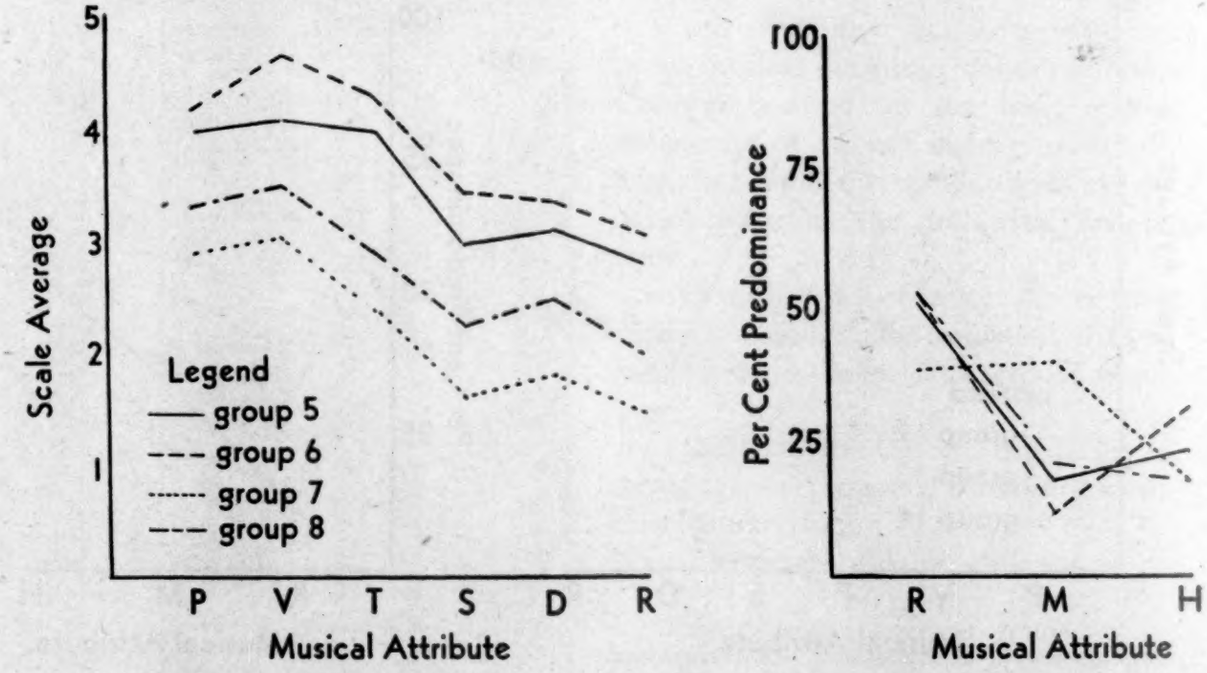

FIG. 2. Musical profiles for Groups 5 (exciting), 6 (very exciting), 7 (dignified), and 8 (kingly)

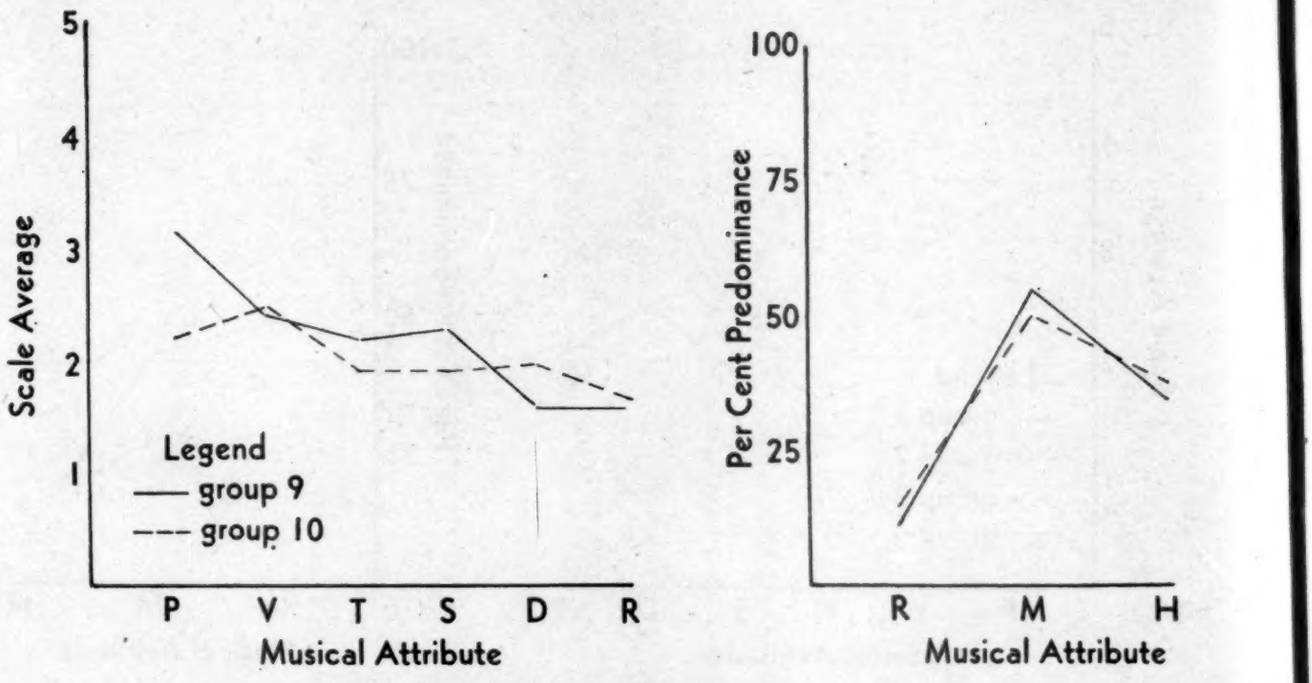

FIG. 3. Musical profiles for Groups 9 (peaceful) and 10 (serious)

ship between *Groups 1* and *2*, therefore, is 29.6; between *Groups 1* and *3* it is 11.1; and so on.

These data are of particular significance in connection with the data of the preceding section (musical characteristics of the adjective-groups). In every instance the degree of relationship between groups is consistent with the degree of similarity between the musical characteristics of the groups in question.

The most closely related groups are *1*

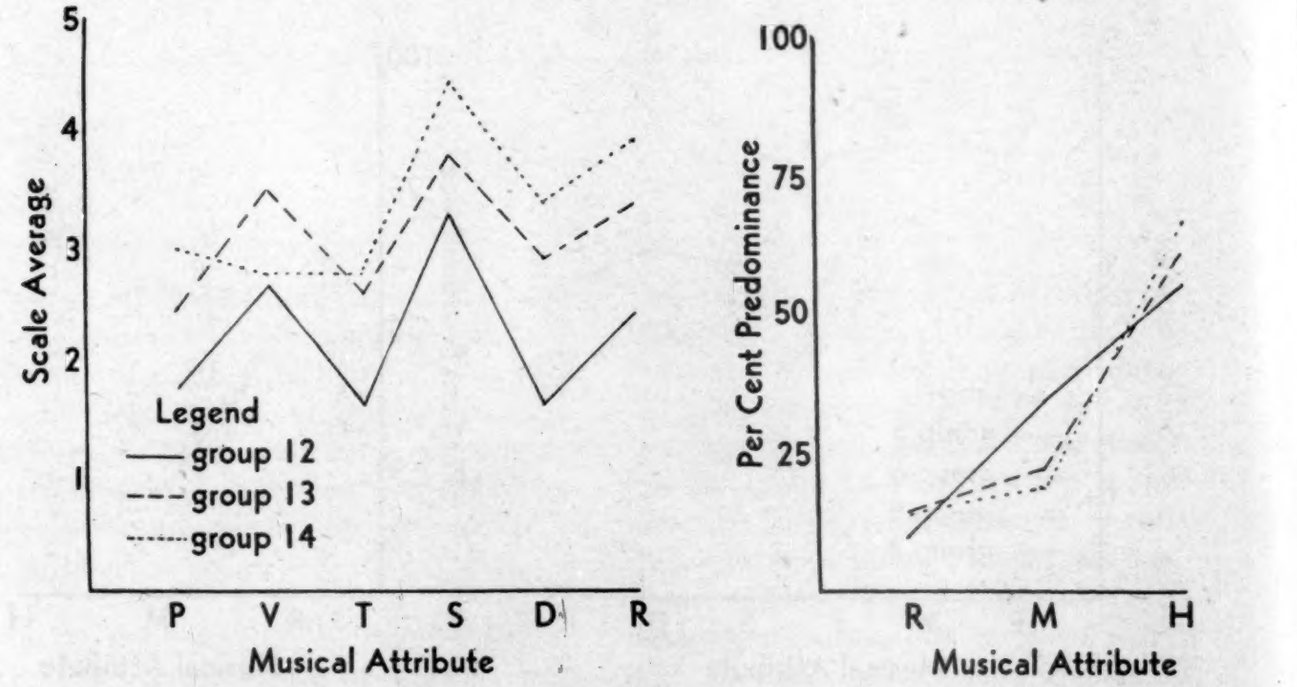

FIG. 4. Musical profiles for Groups 12 (sad), 13 (tragic) and 14 (mysterious)

TABLE 6

Per Cent Interrelationship of Adjective-Groups

Adjective-Group	Adjective-Group 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Happy		29.6	11.1	14.8			11.1	7.4	18.5						7.4
2. Mischievous	41.7		25.0	8.4											
3. Amusing	36.4	45.5		18.2											
4. Very happy	21.0	31.6	5.3		31.6										4.0
5. Exciting	8.0	8.0	8.0	32.0		32.0		8.0							
6. Very exciting				5.6	72.2						5.6	16.7			
7. Dignified	8.7							47.8	17.4	21.7					4 4
8. Kingly	5.0			5.5	25.0		45.0			10.0					10.0
9. Peaceful	8.3						20.8	8.3		37.5		12.5			12.5
10. Serious							26.9	15.4	19.2		11.6	19.2			7.7
11. Pleading										25.0		33.3	33.3		8.3
12. Sad									5.0	15.0	20.0		40.0	15.0	5.0
13. Tragic					11.1	5.6				5.6	11.1	44.4		22.2	
14. Mysterious	10.0				15.0					5.0		20.0	25.0		25.0
15. Unclassified	33.3			8.3			16.7					8.3	8.3		

(happy), 2 (playful), 3 (amusing), and *4 (exuberant); 5 (exciting),* and *6 (sensational); 7 (dignified)* and *8 (kingly); 9 (peaceful)* and *10 (serious); 12 (sad)* and *13 (tragic)* and *14 (mysterious*). The correspondence between relationship of meanings and similarity of musical attributes is best illustrated by the musical profiles[19] for the different meanings.

Profiles are shown for *Groups 1, 2, 3,* and *4* (all closely related) in figure 1; for *Groups 5* and *6* and *Groups 7* and *8* in figure 2; for *Groups 9* and *10* in figure 3; and for *Groups 12, 13,* and *14* in figure 4.

The most important observation from the figures is the similarity of the *shapes* of the profiles for related meanings. This indicates that the entire pattern of musical attributes is important for determining meanings, that musical meanings are not determined by two or three unrelated factors but are determined by constant interrelationships or patterns of the whole array of musical attributes.

SUMMARY

A test of ability to discriminate between musical meanings was devised and was standardized on the basis of the judgments of twenty expert musicians. The data from the standardization procedure warrant the following conclusions:

1. Certain types of music can be consistently classified by musicians into at least fifteen different categories of meanings.

2. Each of the fifteen categories of meanings is represented by definite musical characteristics.

3. There are reliable differences be-

[19] Musical profiles represent graphically the average musical characteristics of the different meanings and the per cent predominance of rhythm, melody, and harmony.

tween the musical characteristics of all fifteen categories of meanings.

4. Interrelation of the different meanings is consistent with the nature of their musical characteristics.

5. Musical meanings are not determined by a few isolated factors but are determined by constant patterns of the entire array of musical attributes.

Chapter IV

Descriptions of Musical Meanings at Different Age-Levels

Although there have been a few experimental studies prior to this investigation in which data have been reported on the nature of musical meanings for adults, there have been no similar experimental data reported on the nature of musical meanings for children. In the preceding sections it has been shown that for musicians certain abstract meanings can be represented by definite musical characteristics and that certain types of music consistently express certain meanings. It is the purpose of this chapter to show to what extent these relationships hold true for children and untrained adults. Analysis of the problem will be made along three lines: (1) relative frequency with which different meanings are attributed to music, (2) musical characteristics of the different meanings, and (3) interrelationship of the different meanings. The constancy of meanings will be further analyzed on the basis of predictions of change in meanings corresponding to controlled changes in the musical characteristics of the standardized selections. In order to simplify the treatment of (1), (2), and (3) above, only general summaries of the data are presented in this monograph. Complete tabular representations can be found in the original manuscript.

EXPERIMENTAL PROCEDURE

The selections in Form A of the test were played for groups of subjects at the sixth-, eighth-, tenth-, and twelfth-grade levels, for college students, and for graduate students. The same portable Victrola which had been used for the standardization of the test was used for all the testing. Each selection was played through twice. On the first hearing, the subjects were requested to check the musical characteristics of the record; on the second hearing subjects indicated the meaning of the selection. In the sixth and eighth grades the writer met with the groups to be tested the day before the actual testing and explained all the terms used in the test. Several selections not used in the test were played as a practice exercise at this time. The testing procedure itself required approximately forty minutes. Subjects indicated on the test form the amount of their musical training.

Subjects were obtained from the Durham Public Schools and from Duke University. The results from 100 subjects at each level are analyzed in this paper. The number actually tested at each level was more than 100, but for purposes of simplifying the analysis of data, results are presented for only this number. The 100 papers at each level were selected at random.

ANALYSIS OF THE EXPERIMENTAL DATA

Types of Meanings Most Frequently Attributed to Music at Different Age-Levels

It is of some importance to know the relative frequency with which different meanings are attributed to music. Do untrained subjects and children attribute some meanings to music more frequently than others? Does music in general express certain meanings for untrained subjects and children more or less frequently than for expert musicians?

The frequency with which each adjective classification was mentioned as a first or second choice was tabulated. Allowing two points for a first choice and one point for a second choice, the total number of points for each adjective-group was determined for each subject. This total number of points will be referred to hereafter as the "weighted frequency of mention."

The average weighted frequency of mention for each adjective-group was determined for each grade-level and for the experts. By considering the experts' analysis as a criterion of the actual amount of weight which *should* be attributed to each adjective-group, it was possible to determine if certain meanings are given a disproportionate weight at any particular level. It was found that in general the meanings mentioned most frequently at one level were also mentioned most frequently at all other levels. A similar conclusion can be drawn in regard to those meanings mentioned least frequently. The reliability of the difference between the average of each grade-level and the expert average was computed. There were no reliable differences for any adjective-group at any level.

It may be concluded, therefore, that the relative frequency with which different meanings are attributed to music is approximately the same for all levels tested.

Musical Characteristics of the Various Meanings for the Different Grade-Levels

The average musical attributes; the relative predominance of rhythm, melody, and harmony; and the range of musical attributes were computed for each adjective-group at each grade-level. The most striking observation from these data is the consistency with which the averages for the musical characteristics agree among the different levels and between the different levels and the experts. This indicates that at all levels tested music can be classified according to abstract meanings with certain prescribed musical characteristics. Consistent reliable differences between the lower and higher levels are found only in the following instances:

1. *Group 3* (*amusing*): increase in regularity of rhythm and dynamics.
2. *Group 4* (*very happy*): increase in volume and tempo.
3. *Group 6* (*very exciting*): increase in pitch, volume, tempo, and harshness of sound.
4. *Group 9* (*peaceful*): decrease in dynamics.
5. *Group 11* (*pleading*): increase in pitch and irregularity of dynamics.

A similar agreement is found in the relative predominance of rhythm, melody, and harmony. There is a consistency between all levels in selecting the outstanding elements for all groups, with the exception of a few slight discrepancies. For example, in *Group 6* (*very exciting*) both rhythm and harmony are outstanding in the sixth grade, whereas only rhythm is outstanding at the other levels. These data are interesting in view of the interpretation most frequently given by musicians with respect to growth in music appreciation. It has generally been believed that younger children respond most frequently to rhythmic elements in the music—that only later do they notice melody, and still later harmony. The results presented here do not bear out this assumption. Rhythm, melody, and harmony were noticed with equal frequency at the sixth grade, whereas rhythm was noticed more frequently at the higher levels. There is

some evidence of growth in ability to appreciate each of the three elements, but there is no consistent change in the frequency with which any of the elements is noticed, nor is there any evidence of a differential rate of growth.

It may be concluded, then, that different meanings have approximately the same average musical characteristics at all levels. Exceptions to this general conclusion will be analyzed later.

Reliability of the Difference between Musical Characteristics of the Adjective-Groups

As has been described for the experts, the reliability of the difference between musical characteristics of the adjective-groups was computed for each grade-level.

At the sixth-grade level there are no reliable differences between *Groups 1* (*happy*), *2* (*mischievous*), *3* (*funny*), and *4* (*very happy*); *Groups 5* (*exciting*), and *6* (*very exciting*); *Groups 7* (*dignified*) and *8* (*kingly*); *Groups 7* (*dignified*) and *9* (*peaceful*); *Groups 10* (*serious*), *11* (*pleading*), and *12* (*sad*); *Groups 13* (*tragic*) and *14* (*mysterious*). *Group 15*[20] is not sufficiently well defined to be classified in any category. It may be concluded, then, that sixth-grade children are unable to discriminate between such fine differences in meaning as are included in the test. There are, however, even at this level certain groups of related meanings between which are found reliable differences. A description of such groups follows:

Group A (test *Groups 1, 2, 3,* and *4—happy, mischievous, amusing, very happy*). Outstanding in rhythm; average pitch, volume, and dynamics; above average consonance and irregularity of rhythm.

Group B (test *Groups 5* and *6—exciting, very exciting*). Outstanding in rhythm and harmony; above average pitch, volume, tempo, and irregularity of dynamics; average sound and regularity of rhythm.

Group C (test *Groups 7* and *9—dignified, peaceful*). Rhythm, melody, and harmony of equal weight; below average in all characteristics.

Group D (test *Group 8—kingly*). Outstanding in rhythm; below average in all characteristics except sound (average).

Group E (test *Groups 13* and *14—tragic, mysterious*). Outstanding in harmony; average, pitch, volume, and irregularity of rhythm; above average harshness of sound and irregularity of dynamics; below average speed.

The general description of meanings given above holds also for the eighth grade, except that *Group 10* (*serious*) is related more closely to *Groups 9* (*peaceful*) and *11* (*pleading*) than to *Group 12* (*sad*).

From the eighth grade up, however, there is a noticeable increase in differentiation of groups, with the exception of *Group 15* (*unclassified*), for which there is no consistant interpretation at any level. For grade ten the only groups between which there are no reliable differences (excluding *Group 15*) are *Groups 2* and *4* (*mischievous* and *very happy*, respectively) and *Groups 11* (*pleading*) and *13* (*tragic*). In the latter

[20] It is interesting to note that *Group 15* (*unclassified*) or (*it doesn't mean any of these to me*) is not represented by any definite group of musical attributes at the levels tested, whereas for the expert musicians it is recognized as expressing a definite type of musical meaning. For the untrained subjects "unclassified" was interpreted as being synonymous with "I don't know," whereas the experts interpreted it as representing what may be called "absolute music," a kind of interpretation which is dependent upon training in the theoretical analysis of musical structure. This explanation accounts for the wide difference in scores between experts and untrained subjects for *Group 15*.

instance there is found a difference in relative predominance of melody and harmony for the two meanings. The twelfth-grade and college levels have reliable differences between all groups except *Groups 2* (*mischievous*) and *3* (*amusing*). Differences are found between all groups at the graduate level except *Groups 7* (*dignified*) and *9* (*peaceful*).

Interrelation of the Adjective-Groups

The number of times each of the adjective-groups was mentioned as a second choice to each of the other adjective-groups (when mentioned as a first choice) was determined. Thus, at the sixth-grade level *Group 1* (*happy*) was mentioned 180 times as a first choice. Of these times, *Group 2* (*mischievous*) was mentioned fifty times as a second choice, *Group 3* (*amusing*) thirty-six times, and so on. The per cent relationship of second choices to first choices was determined in the same manner as for the experts (described in the preceding chapter).

The most significant observation from the data is the consistency between different grade-levels in the interrelationship of the various meanings. For example, *Group 2* (*mischievous*) is most closely related to *Group 1* (*happy*) at all levels, *Group 5* (*exciting*) is most closely related to *Group 6* (*very exciting*), *Group 7* (*dignified*) is most closely related to *Group 8* (*kingly*), and so on. These data further indicate that there are constant factors operating to determine the nature of musical meanings at all levels tested. They also furnish evidence that the children at the lower levels were not seriously handicapped by vocabulary difficulties. If the adjectives themselves were not understood, there could not have been such a consistent relationship between musical meanings.

PREDICTION OF CHANGES IN MEANING

If the meaning of a musical selection is determined by certain known musical characteristics of that selection, it should be possible to change the meaning by changing the musical characteristics. Furthermore, it should be possible to predict what this meaning will be. An experiment was performed to determine the validity of this assumption. Although it was impossible to change such elements in the music as consonance, dynamics, rhythm, melody, and harmony, it was possible to regulate the tempo and volume. According to the experts' judgments, the following differences between groups are most dependent upon differences in these attributes:

1. *Group 1* (*happy*) is slower and not so loud as *Group 4* (*very happy*)—not so loud as *Group 8* (*kingly*).

2. *Group 2* (*mischievous*) is faster than *Group 3* (*amusing*), not so loud as *Groups 4* (*very happy*) and *5* (*exciting*).

3. *Group 3* (*amusing*) is not so loud as *Group 5* (*exciting*).

4. *Group 4* (*very happy*) is louder and faster than *Group 1* (*happy*), louder than *Group 2* (*mischievous*), not so loud as *Groups 5* (*exciting*) and *6* (*very exciting*).

5. *Group 5* (*exciting*) is louder than *Groups 2* (*mischievous*) and *3* (*amusing*), louder and faster than *Groups 8* (*kingly*) and *15* (*unclassified*), and not so loud as *Group 6* (*very exciting*).

6. *Group 6* (*very exciting*) is louder than *Group 5* (*exciting*).

7. *Group 7* (*dignified*) is not so loud as *Group 8* (*kingly*), is louder than *Group 9* (*peaceful*).

8. *Group 8* (*kingly*) is louder than *Groups 1* (*happy*) and *7* (*dignified*), not so loud or fast as *Group 5* (*exciting*).

9. *Group 9* (*peaceful*) is not so loud

TABLE 7

Changes Made in the Musical Characteristics of Different Selections

Adjective-Group	Selection*	Change Made	Predicted Change in Meaning†
1. Happy	A	Increase in V and T	Group *4*
	B	Increase in V	Group *8*
2. Mischievous	A	Decrease in T	Group *3*
		Increase in V	Groups *4* and *5*
3. Amusing	A	Increase in V and T	Group *5*
4. Very happy	A	Decrease in V and T	Group *1*
	B	Increase in V	Groups *5* and *6*
5. Exciting	A	Decrease in V	Group *2*
	B	Increase in V	Group *6*
6. Very Exciting	A	Decrease in V	Group *5*
7. Dignified	A	Decrease in V	Group *9*
	B	Increase in V	Group *8*
8. Kingley	A	Increase in V and T	Group *5*
	B	Decrease in V	Groups *1* and *7*
9. Peaceful	A	Increase in V	Groups *7* and *10*
10. Serious	A	Decrease in V	Group *9*
12. Sad	A	Increase in V and T	Group *13*
13. Tragic	A	Decrease in V	Groups *12, 14*
	B	Increase in V and T	Group *6*
14. Mysterious	A	Increase in V and T	Group *6*

* Selection "A" refers to selections from Form A of the test; selection "B" refers to selections from Form B of the test.

† Adjective-groups listed in the last column represent the meanings which should result from the changes made in the standardized selection. Thus, increasing the tempo and volume of a selection characterized as ***happy*** (*Group 1*) should change the meaning of this selection to ***very happy*** (*Group 4*).

as *Group 10* (*serious*) and *Group 7* (*dignified*).

10. *Group 10* (*serious*) is not so loud nor so fast as *Group 8* (*kingly*), is louder than *Group 9* (*peaceful*).

11. *Group 12* (*sad*) is not so loud nor so fast as *Group 13* (*tragic*).

12. *Group 13* (*sad*) is not so loud nor so fast as *Group 6* (*very exciting*), is louder than *Group 14* (*mysterious*).

13. *Group 14* (*mysterious*) is not so loud nor so fast as *Group 6* (*very exciting*), not so loud as *Group 13* (*tragic*).

An attempt was made to change the selections in each group so as to make them more similar to the closely related groups mentioned above. It was predicted that the meaning would change in accordance with the nature of the musical change. The changes made are shown in Table 7.

Forty college students were tested on the selections as originally standardized and were tested approximately a week later on the same selections changed as described. Table 8 shows a comparison of weightings (two points for a first choice and one point for a second choice) of the different meanings for the original and revised selections. The first rows of the table are read as follows: Selection A (the letters stand for the selection,

TABLE 8

Changes in Meaning Corresponding to Controlled Changes in the Musical Characteristics of the Selections

Adjective Group	Selection	Adjective-Group														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	A^1	43	25	4	17	1		12	5	5	2					
	A^2	*26*	22	8	*33*	10	2		14							4
	B^1	45	20	3	18	2		7	2	9	1					3
	B^2	*22*	21	8	20	7	3	10	*18*	1						10
2	A^1	22	49	8	15	9	2	3		1				1		2
	A^2	20	*33*	*20*	3	4	1	6	5	3	1				15	4
	B^1	29	40	17	20	6		1	1	2						
	B^2	8	*29*	15	*36*	*20*	2		3			1	1		2	2
3	A^1	21	34	37	10	3		1	1	1					6	4
	A^2	6	31	*25*	18	*20*	3							5	2	10
4	A^1	19	23	10	34	14	2	1	5		1	2				5
	A^2	*26*	20	15	*10*	2		11	8	9	3				2	6
	B^1	22	21	7	32	16	2	2	6		2	1		1	1	6
	B^2	10	8	8	*29*	*36*	*15*		7			1		2		4
5	A^1	5	7		16	40	21	7	14		5	1		1		3
	A^2	15	*13*		30	*21*	3	5	24		6					
	B^1	3	5	1	19	45	25	4	12		2	1				
	B^2			2	17	*38*	*43*	1	10							
6	A^1	1	2	4	8	40	54							9	2	
	A^2		2	2	7	*51*	*30*					3		12	13	
7	A^1	5			8			44	36		21		1			3
	A^2	4	1		1			*40*	27	*26*	17		3			1
	B^1	6			1			46	40	6	13		2			6
	B^2	3	2	2	5	6		*38*	*48*		10	1				5
8	A^1	5	3	1	11	20	1	29	41		4	1	1	1		2
	B^2	2	1	1	10	*42*	27	8	*25*		2			1		1
	B^1	3	2		10	25	6	27	38		5		2			2
	B^2	*8*	1		7	10		*34*	*30*	4	9		3		1	3
9	A^1	3			1			23	10	55	19	3	2	1	2	
	A^2	2		4	3	3		*30*	13	*19*	*33*	7	3	3		
10	A^1							13	10	36	48	7	5			
	A^2							16	5	*44*	*35*	3	8		2	4
12	A^1							8	3	5	18	5	54	22	5	
	A^2					5	1				21	7	*30*	*40*	8	
13	A^1			1		7	2	13	14		10	8	22	33	10	
	A^2							16	7	8	15	2	*38*	*19*	15	
	B^2					8	4	10	8	1	12	15	20	30	12	
	B^2					29	20	4	5		5	8	10	*29*	10	
14	A^1		3	1		5	2	2	2	4	5	3	20	17	49	7
	A^2		1	3		30	*27*						12	16	*30*	1

the exponents for original (1) or revised (2) forms) in its original form was given a weighting of 43 for *Group 1*, 25 for *Group 2*, and so on. This same selection, when increased in volume and tempo (nature of change determined from Table 7), was given a weighting of only 26 for *Group 1*, and so on. The most noticeable change in meaning for Selection B was in *Group 8*, which is consistent with the prediction. Changes which were predicted are italicized in the table. For example, *Group 1* differs from *Group 4* chiefly in tempo and volume. When selection A^1, *Group 1*, was played in the original form, a weighting of 43 was given *Group 1* and a weighting of 17 was given *Group 4*. When the selection was increased in tempo and volume, however, the weighting for *Group 1* dropped to 26 and that for *Group 4* increased to 33. These changes are italicized. In every instance, change in the musical elements produced in the subjects' characterizations of the record a change which was consistent with the prediction.

CHAPTER SUMMARY

The meaning of music at different age-levels was analyzed from the standpoint of (1) the relative frequency with which different meanings are attributed to music, (2) the musical characteristics of the different meanings (average pitch, volume, etc.; relative predominance of rhythm, melody, and harmony; and reliability of the differences between musical characteristics of different meanings), and (3) the interrelationship of the various meanings. The constancy of musical meanings has been further analyzed at the college level by determining the validity of predictions in change of meanings corresponding to controlled changes in the musical attributes of the selections.

The data from all the different analyses support the conclusion that the meaning of music is not a fortuitous subjective variable—that there are constant factors in the music itself which determine the nature of individual interpretation at all levels tested. More specific conclusions are:

1. The relative frequency with which certain meanings are attributed to music is approximately the same for all age-levels tested as for the expert musicians.

2. There is a consistency between musical meanings and the musical attributes which represent them for all levels tested. Reliable differences are found between the musical attributes of most meanings at all levels. In general, abstract meanings have approximately the same musical characteristics for untrained adults and for children as for expert musicians.

3. There is a consistency between all levels in the interrelationship of the different musical meanings.

4. Controlled changes in the musical characteristics of a selection produce predictable changes in the meanings attributed to the selection.

Chapter V

Growth of Ability to Discriminate Between Musical Meanings

In chapters III and IV it was shown that there is a definite relationship between certain patterns of musical attributes and the type of meaning which is attributed to music. In the present chapter an analysis is made of growth in ability to recognize and appreciate these relationships. Analysis of growth is made from the standpoint of the following:

1. The growth curve of ability to discriminate between musical meanings.
2. Growth in ability to recognize individual meanings.
3. Growth in appreciation of the musical characteristics of the different meanings.
4. Growth in recognition of the interrelationship of meanings.

In the original manuscript the analyses listed above were based on data obtained from 100 subjects at each grade-level. In the present monograph they are based on the original data plus data from 100 additional subjects tested after the completion of the original study. Likewise included in the present monograph (Chapter VI), but not in the original study, are analyses of the general nature of the ability, based on data from the 100 new subjects, which were not contained in the original study.

EVIDENCE OF GROWTH

The Growth Curve of Ability to Discriminate between Musical Meanings

Individual papers were scored on the basis of the weightings derived in the standardization of the test. The data for each grade-level and for the experts are presented in Table 9.[21] The first row of the table is read as follows: At the sixth-grade level scores ranged from 28 to 141; the group average was 83.85; the standard deviation of the distribution of scores was 23.16; the standard deviation of the average was 1.64. The final column of the table is read as follows: the critical ratio between the eighth- and sixth-grade levels was 3.23; it was 4.55 between the tenth- and eighth-grade levels, etc. These data indicate that there are reliable differences between the averages for each level and each higher level with the exception of the college and graduate levels (C.R. is .64).

The growth curve for ability to discriminate between musical meanings is shown in Figure 5. Care must be taken in interpreting the curve to allow for the factor of selection (caused by certain children leaving school, etc.), which is operating to some extent from level to level. Available data (presented in the original study) indicate, however, that compensation for this factor would not appreciably alter the shape of the curve. It may be concluded that growth in musical understanding (as measured by this test) proceeds at a fairly constant rate from the sixth-grade through the college levels. This continuity of growth is particularly striking in contrast to curves obtained from various tests of the Seashore battery, which represent growth in musical ability as reaching its maturity, in most instances, around the tenth-grade level.

[21] The data presented in Table 9 may be considered as tentative grade-level norms for the test.

Growth in Ability to Recognize Individual Meanings

Although the growth curve shows the differences between averages for different grade-levels, it does not in any way indicate the specific factors which contribute to such growth. In order to determine what factors contribute to the gross results, a comparison was made between the scores obtained by each grade-level on each of the individual types of meaning.

For any one meaning there were several selections in the test which to some extent expressed this meaning. For example, although a given group, say *happy, glad,* is the best choice for only one of the fifteen records, it is also allowed varying degrees of credit when selected for seven other records. The total amount of credit (score) obtained by each subject for each of the adjective-groups (all records included) was computed. From these scores the average score for all subjects at a given level was determined for each adjective-group. Thus, for *happy, glad,* the average total number of points obtained at the sixth-grade level for all the times this group was mentioned was 9.41 per person.[22]

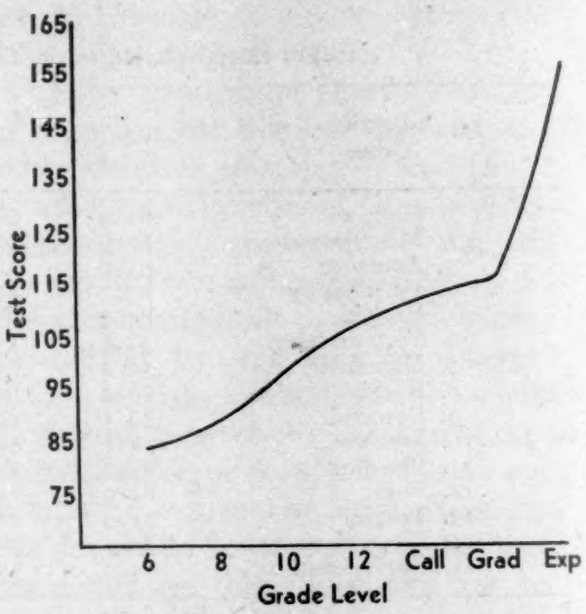

FIG. 5. Growth curve of ability to discriminate between musical meanings.

It was possible to make more points on some meanings than on others. For example, there were eight selections for which *Group I* (*happy*) was allowed some credit, whereas *Group II* (*pleading*) was allowed credit on only two. In order to

[22] Records in the test for which some credit was allowed for this answer are: Record 1: (selection of *happy, glad* as first choice allowed nine points credit, as a second choice six points); Record 2: (first choice five points, second choice four points); Record 3: (first choice three points, second choice two points); Record 4: (first choice four points, second choice four points); Record 5: (first choice two points, second choice two points); Record 8: (first choice two points, second choice two points); Record 9: (first choice two points, second choice two points); Record 15: (first choice four points, second choice three points).

TABLE 9

Ranges, Averages, Standard Deviations, and Critical Ratios of Scores and Differences for Each Grade-Level and for the Experts

Grade Level	Number of Subjects	Range*	Average	S.D. Dist.	S.D. Av.	C.R. of Diff.
Sixth	200	28–141	83.85	23.16	1.64	
Eighth	200	33–143	90.76	19.32	1.37	3.23
Tenth	200	49–145	99.32	18.36	1.29	4.55
Twelfth	200	65–164	107.43	19.01	1.34	4.36
College	200	49–173	114.35	22.13	1.57	3.35
Graduate	100	59–165	116.23	24.83	2.48	0.64
Expert	20	138–180	158.80	14.03	3.10	10.72

* Highest possible score: 182.0.

TABLE 10

Per Cent Relation to the Criterion of Each Group Average

Adjective-Group	Level					
	Sixth	Eighth	Tenth	Twelfth	College	Graduate
1. Happy	64.9	64.4	65.1	73.6	68.7	71.8
2. Mischievous	65.9	68.0	76.7	76.9	82.0	81.3
3. Amusing	33.9	42.1	51.0	62.9	67.5	70.7
4. Very happy	25.2	27.6	45.4	48.5	60.8	60.0
5. Exciting	64.9	64.4	68.1	78.1	87.8	85.5
6. Very exciting	33.1	37.5	74.2	94.8	96.9	99.8
7. Dignified	59.4	61.8	70.4	77.9	79.0	77.8
8. Kingly	60.5	63.1	68.9	80.0	80.4	79.6
9. Peaceful	56.1	59.3	64.3	69.3	80.1	77.9
10. Serious	54.1	54.5	59.9	70.8	70.5	73.2
11. Pleading	25.7	28.1	47.2	44.6	62.9	65.9
12. Sad	60.0	61.8	68.1	74.6	80.9	81.8
13. Tragic	21.9	28.1	52.5	55.8	60.9	63.2
14. Mysterious	33.1	35.5	32.9	47.1	54.5	55.9
15. Unclassified	13.3	15.5	16.5	30.2	41.2	38.9

equate the various groups for purposes of comparison, each average was converted into the per cent relationship to the criterion (experts' analysis). These data are presented in Table 10, which is read as follows: For *Group 1* sixth-grade children, on an average, make 64.6 per cent as many points as do the experts; eighth-grade children do 64.4 per cent as well as the experts on this group; tenth-grade children do 65.1 per cent as well; and so on.

It is apparent from the data that some meanings are considerably more difficult to recognize than others. Table 10 shows that sixth-grade children do only 13.3 per cent as well as the experts on *Group 15* (*unclassified*), whereas they do 64.9 per cent as well for *Group 5* (*exciting*). The most difficult groups for all levels are *Groups 4* (*very happy*), *11* (*pleading*), *13* (*tragic*), *14* (*mysterious*), and *15* (*unclassified*). The groups most easily recognized at all levels are *Groups 2* (*mischievous*), *5* (*exciting*), *7* (*dignified*), *8* (*kingly*), and *12* (*sad*).

These data indicate, further, that there were only small, if any, vocabulary difficulties for specific meanings. There were only three adjective-groups which the writer felt might present language difficulties. These were *Groups 7* (*dignified*), *8* (*kingly*), and *10* (*serious*). If the children had not understood the meaning of these words the score for these groups would have been lower than for other groups. Such was not the case; the scores for all of these groups were higher than scores for most other groups.

With the exception of *Group 1* (*happy*) there is considerable growth in the ability to distinguish each of the adjective-groups. Children in the sixth grade apparently are able to recognize happy music approximately as well as subjects at any other level. Growth is most noticeable in *Groups 3* (*amusing*), *4* (*very happy*), *6* (*very exciting*), *11* (*pleading*), *13* (*tragic*), and *15* (*unclassified*). The differential rates of growth for the different meanings indicates that there are individual factors determining growth in each of the different meanings. That is, growth in each of the different meanings is not determined entirely by a single factor of general musical understanding. Rather, musical understanding is but the total of all the individual

factors which go to make up comprehension of individual meanings.

Growth in Appreciation of the Musical Characteristics of the Different Meanings

It has been indicated in Chapter IV that the musical characteristics of the different meanings were approximately the same for all levels. Consistent reliable differences between the lower and higher levels were found only in the following instances:

1. *Group 3* (*amusing*): increase in irregularity of rhythm and dynamics.
2. *Group 4* (*very happy*): increase in volume and tempo.
3. *Group 6* (*very exciting*): increase in pitch, volume, speed, and sound.
4. *Group 9* (*peaceful*): decrease in dynamics.
5. *Group 11* (*pleading*): increase in pitch and irregularity of dynamics.

Although there are few reliable differences between the averages at different levels, certain trends of growth can be observed. A summary of all the most noticeable evidences of growth follows:

The average musical characteristics of the different meanings are approximately the same for all levels, but there is a noticeable difference in the variability of the ratings. Almost without exception there is a decrease in standard deviation at the higher levels. For example, although the average pitch for *Group 1* is approximately 3.3 at all levels, the standard deviation decreases from 1.02 at the sixth grade to .58 for the experts; volume decreases from .85 to .57; tempo decreases from .86 to .56; and so on. The consistency between levels in the average musical characteristics of the different meanings indicates that there are constant factors in the music itself which determine the meanings; the decrease in standard deviations indicates growth in recognition of these factors.

Another evidence of growth in ability to recognize the musical characteristics of the different meanings is seen in the increase in the number of reliable differences between characteristics of the various groups. It was shown in Chapter IV that sixth-grade children could differentiate reliably between only six different groups, whereas graduate students could distinguish between the characteristics of thirteen different groups. The data also furnish an explanation of the fact that discrimination between some meanings is more difficult than discrimination between others. It was shown in the analysis of relative difficulty of the different meanings that at the lower levels *Groups 4* (*very happy*), *6* (*very exciting*), *11* (*pleading*), *13* (*tragic*), and *15* (*unclassified*) were the most difficult to recognize. In each instance (with the exception of *Group 15*) the meaning is dependent upon a recognition of intense dynamic factors in the representative music. It is apparent that the subjects did not recognize this factor as of particular importance in the meanings. *Group 4* is most clearly related to *Groups 1* (*happy*), *2* (*playful*), and *3* (*amusing*) at the sixth- and eighth-grade levels; it is most closely related to *Group 5* (*exciting*) at the higher levels. At the lower levels the musical characteristics for *Group 6* (*very exciting*) are no different from those of *Group 5* (*exciting*), and the characteristics for neither group are as intense (fast, loud, etc.) as for the higher levels. At the lower levels, *Group 11* (*pleading*) is most similar to *Groups 10* (*serious*) and *12* (*sad*). It is apparent that children do not recognize at all the intensity of feeling in the word "pleading," even though they do know the verbal meaning of the word. The same may be said for

Group 13 (*tragic*). Although the average pitch, volume, etc., for this group are approximately the same for all levels, the standard deviations of the different characteristics at the lower levels are so large that it can only be concluded that children do not recognize the intensity of feeling involved in this word.

The data for average range of musical characteristics for the different meanings were not presented in Chapter IV. There were so few indications of range of choice at the lower levels that the averages would not be reliable. There was a noticeable increase at the higher levels, however, in the number of instances in which ranges were indicated. This suggests, and the observation is supported by data from the preliminary investigation, that at the lower levels the subjects determine their ratings on the basis of only a fragment of the selection, whereas at the higher levels the judgment is made on the basis of the entire selection. This ability to withold judgment and to rate a selection on the basis of isolated fragments is a further evidence of growth.

The most noticeable evidences of growth in appreciation of the musical characteristics of the different meanings are summarized below for each group.

Group 1 (*happy, glad*): Growth indicated particularly in the decrease of standard deviation in pitch, sound, and dynamics and in the decrease of average for irregularity of dynamics and for rhythm. Growth in the direction of recognizing *happy, glad* as being more regular in rhythm and dynamics than the other more closely related groups.

Group 2 (*playful, mischievous*): Except for increase in tempo and a slight increase in pitch, averages for all levels are approximately the same. Chief evidences of growth are the decrease of standard deviations for volume, tempo, and dynamics. Slight increase in the predominance of rhythm.

Group 3 (*amusing, funny*): Growth indicated by a reliable increase in irregularity of rhythm and dynamics and corresponding decrease in standard deviations.

Group 4 (*very happy, exuberant*): Reliable increase at the higher levels in volume and tempo and consistent decrease in standard deviations. Increase in predominance of rhythm.

Group 5 (*exciting, stirring*): Increase of pitch, volume, and tempo the only consistent evidence of growth for this group. Slight increase in the predominance of rhythm.

Group 6 (*very exciting, sensational*): A reliable increase in pitch, volume, tempo, and harshness of sound. Slight increase in predominance of rhythm.

Group 7 (*dignified, stately*): Growth less clearly indicated than in the other groups. Slight decrease in irregularity of dynamics and corresponding decrease in standard deviation.

Group 8 (*kingly, pompous*): Slight increase in volume and decrease in irregularity of rhythm. Decrease in standard deviations for dynamics and rhythm. Increase in predominance of rhythm.

Group 9 (*peaceful calm*): Decrease in irregularity of rhythm and dynamics and corresponding decrease in standard deviations.

Group 10 (*serious, reverent*): No marked evidence of growth. Slight decrease in harshness of sound, irregularity of rhythm and irregularity of dynamics. Decrease in standard deviation for regularity of rhythm. Increase in predominance of melody.

Group 11 (*pleading, passionate*): Growth largely represented by increase in the number of instances in which wide

ranges of pitch and volume are indicated. Increase in average pitch, volume, and irregularity of dynamics. Increase in predominance of melody.

Group 12 (*sad, mournful*): Growth evidenced chiefly in decrease of standard deviations for pitch, sound, and dynamics. Increase in predominance of rhythm.

Group 13 (*tragic, desolate*): Decrease in standard deviations for pitch, volume, tempo, sound, and dynamics. Increase in number of instances in which wide ranges of pitch and volume are indicated Increase in predominance of harmony.

Group 14 (*mysterious, strange*): Decrease of standard deviations for volume, tempo, and sound. Increase in predominance of harmony.

Group 15 (*unclassified*): Evidence of growth limited to decrease in size of all standard deviations. Standard deviations for all levels, however, are much larger than for the experts.

The most important observation in connection with the analysis of growth in appreciation of the musical characteristics of the different meanings is that in every instance (with the exception of *Group 15*) where there is a noticeable change from level to level the change is in the direction of the experts' judgments. This lends further credence to the conclusion that there are constant factors underlying the meaning of music.

Growth in Recognition of the Interrelationship of Meanings

As has been indicated previously, there is a general consistency between the different levels in the interrelationship of the different meanings. There are, however, some evidences of growth. For example, there is a decrease in relationship between *Groups 3* (*amusing*) and *1* (*happy*), *1* and *4* (*very happy*), *9* (*peaceful*) and *8* (*kingly*), *11* (*pleading*) and *9*, *12* (*sad*) and *10* (*serious*); there is an increase in relationship between *Groups 3* and *2* (*playful*), *5* (*exciting*) and *4* (*very happy*), *10* and *9*, *13* (*tragic*) and *11*, *13*, and *12*. In every instance the increase or decrease in relationship is consistent with the musical characteristics of the different adjective-groups as described in Chapter IV. For example, *Group 4* (*very happy*) is more closely related to *Group 5* (*exciting*) at the higher levels than to *Group 1* (*happy*). This is consistent with the observation that the musical characteristics of *Group 4* more closely approximate those of *Group 5* than those of *Group 1* for the higher levels, whereas the opposite is true for the lower levels. Likewise, in every instance of consistent change in per cent relationship the change is in the direction of the expert judgments.

CHAPTER SUMMARY

Growth in ability to discriminate between musical meanings has been analyzed from the standpoint of (1) general growth curves, (2) relative difficulty of discrimination between the different types of meanings, (3) recognition of the musical characteristics of the different meanings, and (4) recognition of the interrelationship between meanings. Specific conclusions drawn from the data are:

(1) There is a consistent growth in ability to discriminate between musical meanings from the sixth-grade through the college levels.

(2) There are differential rates of growth for comprehension of the different meanings. Growth is most marked in those types of meanings which are dependent upon intense dynamic factors in the representative music.

(3) Change in the musical characteristics of the different meanings from level to level is in every instance in the direc-

tion of the experts' judgments. Changes from level to level are noticed in (a) average musical characteristics of the different meanings; (b) standard deviations of the ratings; (c) relative predominance of rhythm, melody, and harmony; (d) reliability of the difference between musical characteristics of the different meanings; and (e) frequency with which ranges of choice are indicated for the different musical characteristics.

(4) Change in the per cent relationship between musical meanings is always in the direction of the experts' judgments.

Chapter VI

Factors Contributing to the Growth of Musical Understanding: Nature of the Ability

The nature of the ability to discriminate between musical meanings, and the factors contributing to its growth, were analyzed from the standpoint of relationship to the following: (1) general intelligence, (2) amount of musical training, (3) "musical ability" as determined from the Seashore tests,[23] (4) enjoyment of music.[24]

Relation of the Ability to General Intelligence

The relationship of general intelligence to the ability of discrimination between musical meanings was determined by comparing the test scores of pupils in "high" sections with those of pupils in "low" sections at the sixth-, eighth-, and twelfth-grade levels. At the college level the correlation between scores on the test and scores on the *American Council Psychological Examination*[25] was determined.

Table 11

Differences in Average Scores Between "High" and "Low" Sections

Grade	Section	N	Average Score	Difference	S.D. of the Difference
Sixth	High	115	86.13		
	Low	85	80.76	5.37	4.37
Eighth	High	112	92.28		
	Low	88	89.96	2.32	4.06
Twelfth*	High	49	107.43		
	Ave.	141	107.61	.18	4.24

* In the twelfth grade the comparison was made between two sections characterized by the principal as "the brightest students in the senior class" and the other sections which were characterized as of "average intelligence."

The data for the sixth, eighth, and twelfth grades are shown in Table 11, which is read as follows: The average of the 115 pupils in the "high" sections of the sixth grade was 86.13; the average of the 85 pupils in the "low" sections was 80.76; the difference between the two is 5.37; the standard deviation of the difference is 4.37.

The only comparison in which coefficients of correlation were obtained was at the college level. Scores on the *American Council Psychological Examination* were available for all students tested. The correlation for 100 college students

[23] *Measures of Musical Ability*, C. H. Stoelting Co., Chicago, Ill.

[24] The first two of these four variables were analyzed in the original study, and tentative data were there presented. In the present monograph the original data are combined with data obtained from additional subjects tested after completion of the original study. Analysis of the last two variables is here made entirely on the basis of data obtained after completion of the original study.

[25] L. L. Thurstone and T. G. Thurstone, *American Council on Education Psychological Examination for High School Graduates and College Freshmen* (Washington: American Council on Education, 1936, 1937, 1938 editions).

TABLE 12

Comparison of Scores for Subjects Having the Most Musical Training With Those for Subjects Having the Least Training

Training	Level					
	Sixth	Eighth	Tenth	Twelfth	College	Graduate
Most	82.83	93.28	109.34	118.66	128.68	127.52
Least	83.91	87.78	96.41	93.76	100.23	101.37
Difference	−1.08	6.50	12.93	25.90	28.45	26.15
S.D. Diff.	4.67	4.21	4.08	3.72	4.17	4.35

between these scores and scores on the music test was .29±.09. This indicates that there is little relation at the college level between general intelligence and ability to discriminate between musical meanings.

The conclusion to be drawn from these data is that ability to discriminate between musical meanings is not closely related to general intelligence at any level tested. The ability must be classified, therefore, as a "special ability." This does not imply, however, that it is dependent upon a single factor; it does indicate that the determining factors are different from those underlying general intelligence.

These data, in addition to throwing some light on the nature of discriminative ability, further indicate that the test used is a valid test of this ability. It is known that vocabulary comprehension is directly related to general intelligence. If scores on the test were influenced to any appreciable extent by the subjects' ability to understand the meanings of the terms used, there would have been reliable differences between the scores of the "high" sections as compared to the scores of the "low" sections. No such differences were found.

Relation of the Ability to Amount of Musical Training

Each subject was asked to state on his test paper the exact amount of his musical training, whether in private lessons or in class. For each level the scores of the sixty subjects having the most musical training were compared with the scores of the sixty having the least. The averages for each level are shown in Table 12.

The data indicate that at the lower levels musical training bears no relation to the ability, whereas at the higher levels it is an important determining factor. The lack of difference between scores at the lower levels does not indicate that children cannot profit from musical training. It may be explained in either of two ways: (1) the younger children have not had enough training to produce a difference in understanding of the kind measured or (2) the type of training has not been such as would increase such understanding. The writer is of the opinion that the latter explanation is the more important one. Most of the children who had received musical training had had private lessons in piano or violin. The methods of instruction used by most private teachers of children are far from conducive to a real understanding and appreciation of the meaning of music.

The data presented above do not necessarily mean, however, that discriminative ability is determined entirely by specialized musical training. Very few of

TABLE 13

Coefficients of Correlation Between Scores on the Test of Musical Understanding and Seashore Tests

Grade-Level	Sense of Pitch	Sense of Intensity	Sense of Rhythm	Tonal Memory	Entire Battery
Sixth	.36 ± .086	.29 ± .092	.30 ± .091	.22 ± .095	.46 ± .079
Eighth	.35 ± .088	.33 ± .089	.31 ± .090	.39 ± .085	.47 ± .078
Tenth	.30 ± .091	.26 ± .092	.33 ± .089	.39 ± .085	.40 ± .084
Twelfth	.22 ± .095	.19 ± .096	.35 ± .088	.42 ± .082	.31 ± .090
College	.21 ± .096	.22 ± .095	.31 ± .090	.43 ± .082	.33 ± .089
Graduate	.23 ± .095	.18 ± .097	.28 ± .092	.40 ± .084	.28 ± .092

the subjects at any level had had specific training in music appreciation of the kind measured by the test. It may be that subjects who continued their training were those who possessed the greater amount of ability originally. That a factor of native ability is operating to some extent is indicated by the fact that some of the highest scores at each level were made by subjects who had had no musical training and by the fact that a few subjects with the greatest amount of musical training were among those making the lowest scores.

Although it may be concluded that musical training does have an important effect upon ability to understand musical meanings, the conclusion does not exclude the probability that there is also an important native constituent.

Relation of Musical Understanding[26] *to Musical Ability as Measured by the Seashore Tests*

The Seashore tests for pitch, intensity, rhythm, and memory were given to 100 subjects at each level who had had the test of musical understanding. Coefficients of correlation were computed between scores on the test of understanding and each of the individual Seashore tests and the battery as a whole. The data are presented in Table 13.

[26] By "understanding" is meant that specific area of understanding which can be measured by the discrimination between abstract meanings.

From the table it is apparent that correlations between musical understanding and the various abilities measured by the Seashore tests are small. In only four instances are the correlations reliable: at the sixth and eighth grades, between understanding and the composite of abilities measured by the entire battery, and at the twelfth-grade and college levels, between understanding and tonal memory.

The lack of relationship is more strikingly illustrated by a comparison of individual scores and averages. For example, the average score on the understanding test of the twenty students making the lowest score on the test of pitch discrimination is slightly, but not reliably, lower at each level than the average score of the twenty students making the highest score on the pitch test. However, at each level there were at least eight students of those who scored extremely low on the pitch test who had above-average scores on the understanding test. In other words, although there is a positive correlation between pitch discrimination and understanding at each level, lack of ability to discriminate between pitches does not necessarily mean that the student will not be able to understand music. In the first place, the meaning of music is not dependent upon such fine differences in pitch as are found in the Seashore Test, and, in the second place,

TABLE 14

Comparison of Scores for Subjects Who Like Music and Those Who Dislike It or Are Disinterested in It

Enjoyment	Grade-Level					
	Sixth	Eighth	Tenth	Twelfth	College	Graduate
Like	86.5	93.7	105.3	110.1	120.3	121.8
Dislike	80.9	87.2	96.4	95.3	110.1	111.7
Difference	5.6	6.5	8.9	14.8	10.2	10.1
S.D. Diff.	3.4	3.7	3.2	3.8	3.7	3.4

pitch is but one of several factors determining meaning.

This same observation can be made in regard to each of the separate tests of the Seashore Battery and to the battery as a whole.

Relation of Musical Understanding to the Enjoyment of Music

Subjects were asked to indicate the degree of their musical enjoyment on the basis of the following scale: (1) like it almost better than anything else, (2) like it very much, (3) like it a little, (4) neither like it nor dislike it, (5) don't like it, (6) dislike it very much, (7) dislike it almost more than anything else. The relation of musical understanding to enjoyment was determined on the basis of a comparison between the average understanding scores obtained by those subjects who expressed a definite liking for music (1 and 2 above) and the average scores of those who expressed a dislike or disinterest in music (4, 5, 6, and 7 above). The data for all grade-levels are presented in Table 14.

From the table it can be seen that there is a reliable, or near-reliable, difference between the musical understanding of subjects who like music and those who dislike it at each level above the eighth grade. Stated in another way, this would indicate that subjects who understand music tend to enjoy it, whereas those who do not understand music tend to dislike it. This observation is particularly significant in view of a comparison which was made between scores on the Seashore tests and enjoyment. There was found to be little or no relationship at all between the two. The conclusion drawn by the writer is that musical understanding is more important for the enjoyment of music, and consequently music appreciation, than is the ability to perceive such minute differences in musical attributes as are found in the Seashore Tests.

Chapter VII

Summary and Conclusions

The purpose of this investigation was to analyze some of the factors connected with the nature of musical meanings at different age-levels. A test of ability to discriminate between various meanings was devised and was standardized on the basis of the judgments of twenty expert musicians. The test was then used for measuring subjects' responses at the sixth-, eighth-, tenth-, and twelfth-grade levels and at the college and graduate levels. Thus, it was possible to compare responses at different age-levels with the experts' analyses.

The data were analyzed from the standpoint of (1) general nature of musical meanings for the experts and for subjects at the different grade-levels and (2) growth in ability to discriminate between the different meanings. The ability to discriminate between meanings was further analyzed in relation to general intelligence, amount of musical training, "musical ability" as determined by the Seashore tests, and enjoyment of music.

The general conclusion to be drawn from the investigation is that musical meanings are determined by constant factors in the music itself—that they are not determined by purely fortuitous subjective variables. Specific data in support of this conclusion are drawn, first, from an analysis of expert musicians' judgments. The data show (1) there are at least fifteen different types of meanings between which experts can discriminate with a high degree of consistency and (2) each of the different meanings can be described in terms of constant musical attributes.

The conclusion is further supported by data obtained from subjects at different grade-levels. Consistency in the interpretation of music is demonstrated in the following instances.

1. Approximately the same relative frequency of meanings is attributed to music at all levels.

2. The musical characteristics of the different meanings are approximately the same for all levels. Reliable differences between the musical characteristics of most meanings are found at all levels.

3. The interrelation of the different meanings is approximately the same for all levels.

4. Changes in meaning may be predicted on the basis of controlled changes in the musical attributes of a selection.

Another type of evidence supporting the constancy of musical meanings is derived from an analysis of growth of meanings. Although the general conclusion emphasizes the similarity of meanings at all levels, there are a few noticeable exceptions. In every instance where exceptions occur, however, there is evidence of growth from level to level in the direction of the experts' judgments. This indicates that even though subjects at the lower levels may not recognize or appreciate the musical factors which determine meaning, these factors themselves are constant and will be recognized when the individual's experience has been more complete. Growth is most noticeable in the following instances:

1. Growth in general ability to discriminate between meanings as shown by the growth curves.

2. Growth in ability to recognize the importance of certain individual musical attributes for determining particular meanings.

3. Growth in appreciation of the interrelationship of certain groups of meanings.

A final type of evidence supporting the constancy of meanings is derived from certain instances of differential rates of growth. In every instance the data can be explained in terms of known predictable factors which are thoroughly consistent with other types of data obtained.

Conclusions as to the relationship of the ability to other factors are:

1. Ability to discriminate between musical meanings is to be classified as a "special ability"—that is, it is not closely related to "general intelligence."

2. There is a definite relationship at the higher levels between amount of musical training and ability to discriminate between musical meanings, no relationship at the lower levels (sixth and eighth grades). The relationship at the higher levels does not exclude the probability, however, of native ability being an important determining factor.

3. There is little or no relationship between musical understanding and "musical ability" as measured by the Seashore tests individually or as a battery. Some subjects at each level making the lowest scores on the Seashore tests can, nevertheless, understand the meaning of music better than the average of their group.

4. Musical understanding is more closely related to musical enjoyment than is "musical ability," as measured by the Seashore tests.

As was stated earlier, musicians agree that appreciation of music is dependent upon musical understanding. It has been generally assumed, however, that musical meanings, with the exception of technical knowledge of music structure, are determined largely by inconstant subjective variables. The data presented in this study indicate that there are certain primary attributes of music, recognizable even by untrained subjects, which bear constant relationships to generally recognized abstract meanings. It is thus possible to teach children and adults to understand and appreciate this aspect of music even though they are unable to master the technical skills of musical performance or the more rigorous detail of music theory and composition. It is the writer's belief that such an approach to music appreciation is fundamental to any course of musical training. The importance to music education of a reliable measure of such musical understanding is readily apparent.[27]

[27] The test as described in this monograph is only a preliminary form. Its final standardization is now being completed by the writer.

FINAL TEST FORM

Name.....; Age.....; Birthday.....; Grade.....

Amount of Music Training (Kind of training you have had, when it was, and how long it lasted. Include private lessons, class lessons, glee clubs, orchestras, bands, etc.)

Selection number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
PITCH																		
Very Low																		
Low																		
Medium																		
High																		
Very High																		
LOUDNESS																		
Very Soft																		
Soft																		
Medium																		
Loud																		
Very Loud																		
SPEED																		
Very Slow																		
Slow																		
Medium																		
Fast																		
Very Fast																		
SOUND																		
Very Pretty																		
Pretty																		
Medium																		
Harsh, Ugly																		
Very Harsh, Ugly																		
DYNAMICS (*Quick changes in loudness*)																		
No Quick Changes																		
Few Quick Changes																		
Medium																		
Many Changes																		
Very Many Changes																		
RHYTHM																		
Very Regular																		
Regular																		
Medium																		
Irregular																		
Very Irregular																		
WHICH DO YOU NOTICE MOST?																		
Rhythm																		
Melody																		
Harmony																		
MEANING																		
First Choice																		
Second Choice																		
Third Choice																		

GROUPS

1. Happy, glad
2. Playful, mischievous
3. Amusing, funny
4. Very happy, exuberant
5. Exciting, stirring, impetuous
6. Very exciting, stormy, sensational
7. Dignified, proud, sedate, stately
8. Kingly, majestic, pompous
9. Peaceful, dreamy, tranquil, calm
10. Serious, reverent, sober, solemn
11. Pleading, yearning, passionate
12. Sad, mournful, gloomy, somber
13. Tragic, desolate
14. Strange, mysterious, afraid, grotesque
15. Unclassified (It doesn't mean any of these to me)